

# UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA AGRONÒMICA I DEL MEDI NATURAL



## IMPACTO DE LA TECNOLOGÍA DE SECUENCIACIÓN MASIVA EN EL CÁNCER DE PULMÓN NO MICROCÍTICO

Grado en Biotecnología-Trabajo Final de Grado

Curso 2015/2016

Alumno: Roberto Tébar Martínez

Directora: Sarai Palanca Suela

Tutora académica: Eloisa Jantus Lewintre



Valencia, 7 de julio de 2016

**Autor/Author:** D. Roberto Tébar Martínez

**Tutor Académico/ Academic Supervisor:** Prof. Dña. Eloisa Jantus Lewintre

**Director-Cotutor/Cotutor:** Dña. Sarai Palanca Suela

**Título:** IMPACTO DE LA TECNOLOGÍA DE SECUENCIACIÓN MASIVA EN EL CÁNCER DE PULMÓN NO MICROCÍTICO.

**Resumen:**

En el cáncer de pulmón no microcítico (CPNM) es necesaria la identificación de biomarcadores que permitan la caracterización molecular del tumor y, con ello, adecuar el manejo de los pacientes. En este sentido, el receptor de crecimiento epidérmico (EGFR) y la cinasa del linfoma anaplásico (ALK) desempeñan un papel relevante como marcadores predictivos de respuesta a la terapia oncológica, en concreto a los inhibidores tirosina cinasa (ITKs). Se han descrito diferentes aproximaciones para su determinación (mutaciones de EGFR y traslocaciones de ALK) siendo la real time-PCR (qPCR) y la FISH, las técnicas más ampliamente utilizadas, respectivamente. Con el desarrollo de las tecnologías de nueva generación, como la secuenciación masiva (*Next Generation Sequencing*; NGS), se abre la posibilidad de analizar simultáneamente diferentes alteraciones oncogénicas *driver* implicadas en la patogénesis del CPNM. Asimismo, esta nueva tecnología permite minimizar algunas complicaciones habituales de la rutina asistencial, como son los altos requerimientos de muestra, la disposición de múltiples tecnologías para abordar los distintos marcadores y un importante consumo de tiempo para proporcionar los resultados al clínico.

**Hipótesis y Objetivos:** La NGS puede ayudar a clasificar el tumor al diagnóstico y guiar a los pacientes con CPNM avanzado en el tratamiento con terapias dianas. Por ello, el objetivo del siguiente trabajo de fin de grado (TFG) será una breve contextualización de la medicina personalizada en el CPNM y la comparación de los métodos empleados actualmente en el laboratorio de diagnóstico frente a nuevas tecnologías desarrolladas, como la NGS.

**Sujetos de estudio:** Se plantea un estudio retrospectivo efectuado sobre muestras seleccionadas de tejido tumoral (FFPE) de pacientes con CPNM avanzado atendidos en el Servicio de Oncología Médica del Hospital Universitario y Politécnico La Fe. Durante el desarrollo del proyecto, se estima el estudio de 40 pacientes. Todos ellos deberán haber manifestado por escrito su consentimiento informado.

**Instrumentalización y determinaciones:**

- Extracción de ácidos nucleicos (ADN y ARN) a partir de material tumoral incluido en parafina (FFPE). Este protocolo de trabajo incluye el desparafinado de la muestra y la obtención y purificación de los ácidos nucleicos. Para ello se utilizarán los kits

*GeneReader DNA FFPE Kit* (QUIAGEN) para ADN y *RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE* (ThermoFisher) para ARN.

- Determinación de mutaciones en el gen *EGFR* mediante *cobas® EGFR Mutation Test v2 (CE-IVD)* de *Roche Molecular* en la plataforma *Cobas® 480z Analyzer*, que permite la detección de mutaciones puntuales, deleciones y/o inserciones en los exones 18, 19, 20 y 21. Los casos positivos serán confirmados por secuenciación directa (*Applied Biosystems® 3130 Genetic Analyzer*) utilizando los protocolos recomendados por el fabricante.
- Secuenciación masiva mediante la tecnología del *IonTorrent* (secuenciación por semiconductores) en la plataforma *Personal Genome Machine™* (PGM™; *Applied Biosystem*). Inicialmente, se requiere la preparación de las librerías, empleándose para ello los paneles *Ampliseq Oncomine Solid Tumor DNA kit* y *Oncomine® Solid Tumour Fusion Transcript kit* (*Thermo Fisher Scientific*) que permiten el análisis de 22 oncogenes y cuatro transcritos de fusión, respectivamente.

**Palabras claves:** biología molecular, cáncer de pulmón, secuenciación masiva, EGFR.

---

**Title:** IMPACT OF THE MASSIVE SEQUENCING TECHNOLOGIES IN NON-SMALL CELL LUNG CANCER

**Summary:**

Identification of biomarkers that allow the molecular characterization of the tumor is necessary in non-small cell lung cancer (NSCLC) in order to provide the patient the most appropriate treatment. In this sense, the epidermal growth factor receptor (EGFR) and the anaplastic lymphoma kinase (ALK) play an important role like predictive biomarkers in response to oncology therapy, in special to tyrosine kinase inhibitors (TKIs). Some approaches have been defined in order to determinate those biomarkers (EGFR mutations and ALK translocations), being real-time PCR (qPCR) and fluorescence *in situ* hybridization (FISH), respectively, the most relevant technologies used in clinic. With the development of new generation technologies like massive sequencing (*Next Generation Sequencing*; NGS), it is possible to identify oncogenic *driver* mutations implicated in the NSCLC pathogenesis. This recent technology can minimize some common complications of routine care, like high sample requirement, diverse technologies used to identify different biomarkers and the long times required to provide results to the clinic

**Hypothesis and objectives:** NGS could help to classify tumor in diagnosis and guide advanced NSCLC patients to appropriate targeted therapies. Thus, the objective of this “Trabajo final de grado” (TFG) is contextualize personalized medicine in NSCLC treatment and compare the habitual methods used in a diagnostic laboratory with new technologies like NGS.

**Study subjects:** A retrospective study in formalin-fixed paraffin embedded (FFPE) tissue samples is proposed. Samples of patients with advanced NSCLC are from Medic Oncology Service of “Hospital Universitario y Politécnico La Fe”. During the development of this project, it is estimated to study 40 patients who must sign the informed consent.

#### Instruments and methods:

- Nucleic acids isolation (DNA and RNA) from FFPE tissue samples. The process includes a previous step of paraffin removal, and isolation and purification of the nucleic acids. DNA will be obtained using *GeneReader DNA FFPE Kit* (QUIAGEN) and RNA with *RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE* (ThermoFisher).
- Detection of mutations in *EGFR* gene with *cobas® EGFR Mutation Test v2 (CE-IVD)* of *Roche Molecular* in the platform Cobas® 480z Analyzer. This kit is able to detect point mutations and indels in exons 18, 19, 20 and 21. Positive samples will be confirmed by capillary sequencing (Applied Biosystems® 3130 Genetic Analyzer) according to fabricant protocols.
- Massive sequencing using IonTorrent technology (sequencing by semiconductors) in the platform *Personal Genome Machine™* (PGM™; Applied Biosystem). Previously, libraries are prepared employing Ampliseq Oncomine Solid Tumor DNA kit and Oncomine® Solid Tumour Fusion Transcript kit (Thermo Fisher Scientific). The first is capable to identify 22 oncogenes and tumor suppressor genes, and the second can detect the expression of four fusion transcripts.

**Key words:** Molecular biology, lung cancer, massive sequencing, EGFR.



Valencia, 7 de Julio de 2016

## Agradecimientos

Quiero aprovechar este espacio para darle las gracias a la Dra. Sarai Palanca por la oportunidad de realizar este TFG en un laboratorio como es el de Biología Molecular del Servicio de Análisis Clínicos del Hospital Universitario y Politécnico La Fe. Agradecer también a la Dra. Rosa Murria el apoyo que me ha dado durante mi estancia de prácticas.

También agradecer a todas aquellas personas que me han apoyado durante estos 4 años: mis compañeros de “la puerta 5”, mi hermano Eduardo, mi padre, mis primos Iris y Alejandro, Maiden, el resto de mi familia, y todos mis compañeros de los Scouts y a mis niños de la Unidad Kimberley.

Pero sobretodo, si de alguien me acuerdo en este trabajo, es de mi madre. Gracias a ella he llegado hasta aquí, aunque este es solo el comienzo de un largo camino. Cada paso que he dado, que doy y que daré será por ti.

## Contenido

|        |  |    |
|--------|--|----|
| 1.     | Introducción.....  | 1  |
| 1.1.   | Cáncer de pulmón .....   | 1  |
| 1.2.   | Biomarcadores, Medicina personalizada y Cáncer .....                         | 1  |
| 1.3.   | Genes implicados en CPNM.....  | 2  |
| 1.3.1. | Receptor del factor de crecimiento epidérmico ( <i>EGFR</i> ) .....          | 2  |
| 1.3.2. | Cinasa del linfoma anaplásico ( <i>ALK</i> ) .....                           | 4  |
| 1.3.3. | Otros oncogenes o genes supresores de tumores .....                          | 4  |
| 1.4.   | Tratamiento con Inhibidores tirosina cinasa (ITK) .....                      | 5  |
| 1.5.   | Detección de biomarcadores.....  | 7  |
| 1.5.1. | Detección de mutaciones en <i>EGFR</i> .....                                 | 7  |
| 1.5.2. | Detección de reordenamientos en <i>ALK</i> .....                             | 7  |
| 1.6.   | Secuenciación masiva .....   | 8  |
| 1.7.   | Secuenciación dirigida en CPNM.....  | 9  |
| 2.     | Objetivos.....   | 10 |
| 3.     | Material y métodos .....   | 10 |
| 3.1.   | Muestras.....  | 10 |
| 3.2.   | Extracción de ácidos nucleicos.....  | 11 |
| 3.2.1. | Extracción de ADN .....  | 11 |
| 3.2.2. | Extracción de ARN .....  | 12 |
| 3.3.   | Cuantificación de ácidos nucleicos.....                                      | 12 |
| 3.4.   | Detección de mutación mediante PCR en tiempo real (plataforma Cobas®).....   | 13 |
| 3.5.   | Detección de mutación mediante PCR convencional y Secuenciación directa..... | 15 |
| 3.6.   | Detección de mutación mediante secuenciación masiva (plataforma PGM®).....   | 16 |
| 3.6.1. | Genoteca de ADN .....  | 16 |
| 3.6.2. | Genoteca de ARN .....  | 17 |
| 3.6.3. | Combinar librerías .....   | 18 |

|        |  |    |
|--------|--|----|
| 3.6.4. | PCR en emulsión .....  | 18 |
| 3.6.5. | Enriquecimiento de la muestra en esferas con PCR clonal.....               | 19 |
| 3.6.6. | Carga del Chip y Secuenciación .....                                       | 19 |
| 3.7.   | Análisis bioinformático .....  | 20 |
| 4.     | Resultados .....   | 20 |
| 4.1.   | Resultados obtenidos mediante PCR en tiempo real (plataforma Cobas®) ..... | 20 |
| 4.2.   | Confirmación de resultados mediante secuenciación directa.....             | 21 |
| 4.3.   | Estudios de secuenciación masiva mediante PGM .....                        | 24 |
| 4.3.1. | Mutaciones en el panel de ADN.....   | 24 |
| 4.3.2. | Traslocaciones en el panel de ARN.....                                     | 27 |
| 4.4.   | Comparación de resultados entre los métodos moleculares empleados .....    | 27 |
| 5.     | Discusión.....   | 29 |
| 5.1.   | Comparación qPCR, secuenciación directa y NGS.....                         | 29 |
| 5.2.   | Impacto de la tecnología de NGS .....                                      | 31 |
| 6.     | Conclusiones .....   | 33 |
| 7.     | Referencias .....  | 35 |
| 7.1.   | Bibliografía .....   | 35 |
| 7.2.   | Recursos Web .....   | 40 |
| 8.     | Anexos .....   | 41 |
| 8.1.   | Anexo I: Muestras procedentes de un mismo paciente.....                    | 41 |
| 8.2.   | Anexo II: chips empleados para la secuenciación mediante PGM.....          | 42 |
| 8.3.   | Anexo III: Mutaciones detectadas mediante secuenciación en el PGM.....     | 43 |
| 8.4.   | Anexo IV: Resultados de la secuenciación de los transcritos.....           | 46 |

## Índice de Figuras

Figura 1 - Rutas de señalización implicadas en el CPNM mediante activación por EGFR

Figura 2 - Ejemplos de estructura química de ITK y su unión a sus dianas

Figura 3 - Deaminación oxidativa de la citosina en muestras FFPE

Figura 4 - Esquema de la carga de una placa de 96 pocillos para el *Cobas 480z Analyzer*

Figura 5 - Gráfico de sectores de los resultados obtenidos mediante *Cobas Analyzer*

Figura 6 - Captura con todas las bandas de los geles realizados para testar las PCRs destinadas a secuenciación directa

Figura 7 - Electroferogramas de las regiones de interés de los exones 18, 19, 20 y 21 de las muestras y un control negativo (wt), obtenidos mediante secuenciación directa

Figura 8 - Distribución de las mutaciones identificadas en *EGFR* a lo largo de todas las muestras de NGS

Figura 9 - *Heatmap* de las mutaciones de las muestras analizadas mediante PGM con el *Oncomine Solid Tumor DNA kit*

Figura 10 - Distribución de las mutaciones localizadas en *KRAS* mediante NGS

Figura 11 - Alineamientos de las lecturas obtenidas de dos de las muestras conflictivas

## Índice de Tablas

Tabla 1 - Mutaciones de *EGFR* detectadas mediante la plataforma Cobas

Tabla 2 - Cebadores empleados para los distintos exones del gen *EGFR* (18, 19, 20 y 21)

Tabla 3 - Comparación de los resultados obtenidos mediante las técnicas de análisis del gen *EGFR*.



## Listado alfabético de las abreviaturas empleadas

ADN – Ácido desoxirribonucleico

ALK – *Anaplastic Lymphoma receptor Kinase*

ARN – Ácido ribonucleico

ATP – Adenosin Trifosfato

BTC – Betacelulina

ct – *circulating tumor*

COLD - *co-amplification at lower denaturation temperature*

CP – Cáncer de pulmón

CPM – Cáncer de pulmón microcítico

CPNM – Cáncer de pulmón no microcítico

DNA – *Deoxyribonucleic acid*

EGF – *Epithelial Growth Factor*

EGFR – *Epidermal Growth Factor Receptor*

EML4 – *Echinoderm Microtubule associated protein-like 4*

EMT- Transición Epitelio Mesenquimal

FFPE – *Formalin-fixed, Paraffin-embedded*

FISH – *Fluorescence In Situ Hybridization*

GTP –Guanidin Trifosfato

HPLC – *High Pressure Liquid Chromatography*

HRM – *High Resolution Melting*

HUyP – Hospital Universitario y Politécnico

ISP – *Ion Sphere Particle*

JAK – *Janus Kinase*

MAP – *Mitogen Activated Protein*

mi - micro

NCI – *National Cancer Institute*

NGS – *Next Generation Sequencing*

NV – No valorable

PDB – *Protein Data Bank*

PGM – *Personal Genome Machine*

PI3K – Fosfoinositol 3 cinasa

PNA – *Peptide Nucleic Acid*

qPCR – *Quantitative Polymerase Chain Reaction*

rcf – *Relative Centrifugal Force*

RNA – *Ribonucleic Acid*

RTK – Receptores tirosina cinasa

SH2 – *Scr Homologus 2*

STAT - *Signal Transducer and Activator of Transcription*

TCGA – *The Cancer Genome Atlas*

TE – Tris-EDTA

TSG – *Tumor Suppressor Gene*

TFG- $\alpha$  - *Transforming Growth Factor  $\alpha$*

SEOM – Sociedad Española de Oncología Médica

WES – *whole exome sequencing*

WGS – *whole genome sequencing*

WHO – World Health Organization

wt – *wild type*

# 1. Introducción

## 1.1. Cáncer de pulmón

La *World Health Organization* (WHO; 2016) define el cáncer como un término genérico que se refiere a todas aquellas enfermedades caracterizadas por crecimientos anormales de células. Inicialmente, en el año 2000 se establecieron un total de seis propiedades características del cáncer (o *hallmarks*), las cuales posteriormente fueron ampliadas a diez, siendo éstas la inestabilidad genómica y capacidad de mutación, metabolismo energético desregularizado, resistencia a la muerte celular, constantes señales proliferativas, evasión de los supresores del crecimiento, escape inmunológico, inmortalidad replicativa, capacidad inflamatoria, capacidad angiogénica, y capacidad invasiva y metastásica (Hanahan y Weinberg, 2011).

Atendiendo a los datos registrados en el año 2014, el cáncer de pulmón (CP) es una enfermedad que supuso el 23% de las muertes por cáncer en hombres y el 13% en mujeres a nivel mundial, representando el cáncer de mayor mortalidad. La incidencia descrita para este tumor es la mayor entre todos los tipos de cáncer alcanzando la cifra del 13%, porcentaje similar tanto para varones como para mujeres (Ferlay *et al.*, 2015). En la población española, la incidencia del CP en el año 2012 fue del 12,39%, por detrás del cáncer de próstata y del cáncer colorrectal, mientras que la mortalidad fue del 20,55%, la segunda mayor, por detrás de los cánceres del intestino grueso (SEOM, 2016).

Se distinguen los siguientes tipos de CP (García-Foncillas *et al.*, 2011):

- Cáncer de células grandes o no microcítico (CPNM): supone entre el 80-85% de los casos. Engloba los tipos histológicos epidermoide, adenocarcinoma y de células grandes.
- Cáncer de células pequeñas o microcítico (CPM): supone entre el 15 y el 20% de los CP encontrados.
- Otros tipos de CP, como los carcinoides, los tumores pleomórficos, los sarcomatoides y los derivados de glándula salivar.

## 1.2. Biomarcadores, Medicina personalizada y Cáncer

El *National Cancer Institute* (NCI) de USA describe como biomarcador a “cualquier molécula biológica localizada en sangre, otros fluidos o cualquier tejido del cuerpo que sea indicio de un proceso normal o anormal, o de una condición o enfermedad; también pueden ser moléculas que reflejen el comportamiento del cuerpo ante un fármaco” (Henry y Hayes, 2012). Cada biomarcador puede aportar información acerca de diferentes aspectos relacionados con todo el proceso oncológico, desde el riesgo a su aparición hasta un control para identificar posibles recaídas. Así, según el *Biomarker definitions working group* (2001) podemos agrupar los biomarcadores relacionados con el cáncer de la siguiente forma:

- Marcadores previos al cáncer, que determinen el riesgo de enfermedad.

- Marcadores diagnósticos que indiquen la presencia del cáncer y qué tipo es.
- Marcadores que complementen la información del diagnóstico, entre los que encontraremos marcadores pronósticos, acerca de las posibilidades de supervivencia; predictivos de respuesta a determinados fármacos; farmacocinéticos, sobre la dosis que el paciente será capaz de soportar; y monitorizadores de la respuesta, sobre como el paciente evoluciona frente al tratamiento.
- Marcadores que permiten evaluar si está sucediendo una recurrencia o no.

El desarrollo de nuevos fármacos biológicos junto con la identificación de biomarcadores ha permitido dar un nuevo enfoque a los protocolos terapéuticos. Tradicionalmente, los agentes quimioterápicos o citotóxicos han sido utilizados para tratar a todos los pacientes con cáncer, sin entrar a distinguir en subtipos. Tras los avances de los últimos años en el entendimiento molecular a nivel genómico, proteico y metabólico, el desarrollo de terapias dirigidas y la identificación de biomarcadores eficaces, ha surgido un nuevo planteamiento en la terapia oncológica: la posibilidad de caracterizar cada tumor mediante biomarcadores, y adjudicar a aquéllos que cumplan ciertos requisitos el tratamiento más eficaz. Esto es lo que pretende la medicina personalizada, también denominada medicina de precisión: que cada paciente reciba el tratamiento más adecuado (Biankin *et al.*, 2015; González de Castro *et al.*, 2015). Inicialmente, la búsqueda de biomarcadores se centraba en la búsqueda de genes que se correspondieran con el transporte, la metabolización y la respuesta a los fármacos. Con una mayor comprensión de los mecanismos moduladores de la expresión génica y del funcionamiento de los tumores, se ha ampliado la búsqueda a marcadores de carácter epigenético (tanto patrones de metilación como expresión de miARNs), así como al estudio de exosomas, y del ADN tumoral circulante (*circulating tumor DNA* (ctDNA)) (Ingelman-Sundberg, 2015).

### 1.3. Genes implicados en CPNM

#### 1.3.1. Receptor del factor de crecimiento epidérmico (EGFR)

El receptor del factor de crecimiento epidérmico (*Epidermal Growth Factor Receptor* (EGFR)), también llamado HER1 o c-erbB1 pertenece a la familia de receptores tirosina cinasa (RTK) llamados c-erbB, integrada por receptores transmembrana de naturaleza similar, entre los que se incluye el mismo EGFR, HER2 (c-erbB2), HER3 (c-erbB3) y HER4 (c-erbB4) (Herbst, 2003). Como cualquier miembro de los RTK, esta glucoproteína transmembrana se compone de (Brambilla y Gazdar, 2009):

- Región extracelular: algunos ligandos conocidos son el Epithelial Growth Factor (EGF), Transforming Growth Factor  $\alpha$  (TFG- $\alpha$ ), la Betacelulina (BTC) o la epigenina (Harris *et al.*, 2003; Herbst, 2003).
- Dominio espaciador transmembrana tipo hélice.
- Dominio intracelular juxtamembrana.

- Dominio Tirosina cinasa: los residuos de tirosina, inicialmente no fosforilados, se encuentran en regiones conservadas que van a dar especificidad de cuál va a ser el ligando que se una cuando esté activo.
- Región reguladora C-terminal

El funcionamiento del receptor es como el de otros receptores tirosina cinasa (TK). La unión del ligando a la molécula monomérica induce la dimerización homomérica o heteromérica con otra subunidad de la familia *erbB*. Así, los cuatro receptores pueden formar distintas combinaciones entre ellos en función del ligando que los reconozca (Yarden y Sliwkowski, 2001). Sólo el homodímero de *erbB*-3 carece actualmente de ligando y señalización conocida.

La conformación adquirida por ambas moléculas va a favorecer un primer proceso de fosforilación cruzada y un segundo proceso de autofosforilación de las tirosinas situadas en regiones conservadas. Esta conformación activa del receptor le va a permitir reclutar moléculas con dominios SH2 (reconocen el dominio tirosina cinasa fosforilado). Este será el punto de partida para la propagación de las cascadas de señalización derivadas de este fenómeno de reclutamiento-activación. Las rutas implicadas, como la de las MAP Cinasas, la ruta de PI3K o la ruta JACK-STAT van a activar genes implicados en funciones que van a favorecer el crecimiento y la proliferación, pero van a impedir la apoptosis, la diferenciación o la adhesión celular (Figura 1). La pérdida de regulación en este receptor va a favorecer la aparición de características tumorales en la célula (Bambrilla y Gazdar, 2009; Scaltriti y Baselga, 2006).

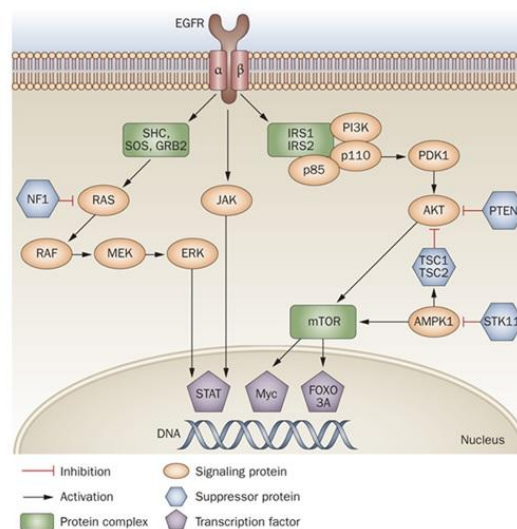


Figura 1: Rutas de señalización implicadas en el CPNM mediante activación del EGFR. La imagen muestra como la unión del ligando favorece la dimerización y el inicio de la cascada de señalización que desencadena la expresión de genes implicados en diversas funciones relacionadas con fenotipos tumorales. Tomado de Harris, T. y McCormick, F. (2010)

Aproximadamente, hasta el 15 % de los casos de CPNM (Gómez *et al.*, 2012) se pueden asociar a la presencia de mutaciones en el dominio tirosina cinasa del receptor EGFR. Existen otros mecanismos que pueden derivar en el incremento de la actividad de este receptor, como la amplificación génica o la sobreexpresión proteica, aunque solo son de interés clínico aquellos causados por mutaciones. Los ensayos clínicos realizados han demostrado que las mutaciones

en el dominio TK son el biomarcador predictivo de respuesta a los tratamientos dirigidos (Lindeman et al., 2013). Las mutaciones más frecuentes, representado casi el 90% de las mutaciones en conjunto son las deleciones *in frame* de 15 y 18 nucleótidos del exón 19 y la mutación puntual c.2573 T>G (p.L858R) localizada en el exón 21. Existen otras menos frecuentes como la p.G719A/C/S del exón 18, la p.S768I del exón 20 o la p.L861Q del exón 21 (García-Foncillas et al., 2011; Gonzalez et al., 2012), que también son evaluadas en clínica. En el exón 19 hay otras deleciones menos frecuentes (de 9, 12, 24 y 27 nucleótidos) e incluso inserciones (de 15 y 18 pb) que también deben ser incluidas en los diagnósticos (Lindeman et al., 2013). En estos casos, se habla de que las células tumorales se vuelven “adictas” a las señales de EGFR. Todos estos biomarcadores, atendiendo a la clasificación inicial, se clasificarían como marcadores diagnósticos por asociarse a la enfermedad y predictivos de respuesta a los Inhibidores Tirosina Kinasa (ITK).

### 1.3.2. Cinasa del linfoma anaplásico (ALK)

La cinasa del linfoma anaplásico (*Anaplastic Lymphoma receptor Kinase (ALK)*) es también un RTK el cual está relacionado con el desarrollo del cerebro. La principal característica en común de los reordenamientos de *ALK* es la obtención de una proteína de fusión en cuya región amino terminal existe un dominio de oligomerización que causa la formación de oligómeros de la proteína de fusión y la autofosforilación de los dominios cinasa de ALK (Soda et al., 2007; Cui et al., 2011). Así, los reordenamientos favorables para un tumor consiguen la pérdida de regulación de ALK y su constante fosforilación, creando una señalización ininterrumpida. En el caso del CPNM, destaca la proteína aberrante con fragmentos del *Echinoderm Microtubule associated protein-like 4 (EML4)*, conocida como EML4-ALK. El mecanismo de formación de tal proteína es debido a una inversión en el brazo corto del cromosoma 2, localizándose *EML4* en la región 2p21 y *ALK* en la 2p23, encontrándose ambos genes en sentidos inversos (Soda et al., 2007). Otros genes implicados en los reordenamientos con *ALK* son *TFG*, *KIF5B* y *KLC1* (Lindeman et al., 2013; Tuononen et al., 2013). Se ha descrito una prevalencia en torno al 4-6% de CPNM portadores de reordenamientos del gen *ALK*.

### 1.3.3. Otros oncogenes o genes supresores de tumores

Además de las mutaciones activantes de *EGFR* o el transcrito de fusión *EML4-ALK*, hay otros genes implicados en la adquisición de un fenotipo tumoral en CPNM, como los que se describen a continuación:

- *KRAS* es un gen que codifica para una GTPasa que continúa la cascada de señalización de RTK, como *EGFR*. Mutaciones en los codones 12 y 13 activan constitutivamente este protooncogen, aunque otros codones de importancia son el 61 y el 59. Esto es conseguido al alterar la actividad GTPasa, por la que *KRAS* pasaría a un estado activo al no poder hidrolizar el GTP. Debido a que el paso a GDP es necesario para la inactivación, el bloqueo de tal proceso impediría la vuelta de *KRAS* a su estado inactivo, causando un exceso de señalización. Se ha llegado a observar una incidencia de entre el 8% y el 24% en CPNM (Luo y Lam, 2013).

- *MET* es un gen implicado en diversas rutas de señalización implicadas en procesos celulares entre los que se encuentran la angiogénesis, la invasividad celular, la transición epitelio mesenquimal (EMT) o la proliferación celular. La desregulación y activación constitutiva de *MET* se consigue mediante diversos mecanismos como la amplificación genómica, sobreexpresión, splicing alternativo o mutaciones activantes (Luo y Lam, 2013).
- *TP53* es un gen supresor de tumores (TSG) que codifica para la proteína p53, la cual, en respuesta a señales es capaz de unirse al ADN activando la expresión de algunos genes o reprimiendo otros. En conjunto, esta regulación promueve el arresto del ciclo celular y la reparación del ADN, la apoptosis, la senescencia o la apoptosis. Además de poseer un dominio de unión ADN, cuenta con una serie de aminoácidos susceptibles de sufrir modificaciones post-traduccionales. Alteraciones en este dominio o estos residuos van a alterar la función de la proteína, impidiendo la actividad antitumoral de p53 (Gibbons *et al.*, 2014).
- Otras alteraciones implicadas en el desarrollo y progresión del CPNM son la presencia de mutaciones en *STK11*, *SMAD4*, *DDR2*, *FGFR3*, *AKT1*, *PIK3CA* o *PTEN*, así como reordenamientos en *ROS1*, *RET* o *NTRK1*.

#### 1.4.Tratamiento con Inhibidores tirosina cinasa (ITK)

Los ITK son moléculas que se unen a las proteínas con dominio TK, inhibiendo su actividad catalítica. Dentro de los pacientes de CPNM, los ITK han cobrado gran relevancia. La función de los ITK es impedir que la cascada de señalización mediada por dominios TK sea continuada, anulando su efecto.

Importantes estudios clínicos en pacientes con CPNM, como el IPASS, con selección por criterios clínicos (Mok *et al.*, 2009), o el Optimal, con selección de pacientes con mutación en *EGFR* (Zhou *et al.*, 2011), mostraron una tasa de respuesta mayor en pacientes tratados con ITK, siendo Gefitinib el fármaco empleado en el primer estudio y Erlotinib el del segundo.

Actualmente existen tres generaciones de ITK desarrollados para tratar pacientes con mutaciones en el gen *EGFR*:

- 1ª generación: En este grupo de agentes terapéuticos se engloban moléculas de unión reversible al RTK portador de mutaciones activadoras del dominio TK. De ellos, en CP destacan los siguientes:
  - Gefitinib (Wakeling *et al.*, 2002): esta 4-anilinoquinazolina actúa como inhibidor competitivo que impide la unión del ATP con K745, un sitio crítico para la unión del ATP (Figura 2).
  - Erlotinib (Stamos *et al.*, 2002): de forma similar al anterior, esta 4-anilinoquinazolina bloquea el sitio de unión del ATP (Figura 2).
  - Crizotinib (Cui *et al.*, 2011): Esta aminopiridina se une a las moléculas de ALK aberrantes, inhibiendo la fosforilación.
- 2ª generación: Las moléculas de esta generación son capaces de unirse irreversible a las moléculas de *EGFR* con mutaciones activadoras del dominio TK, lo que permite que el

tratamiento sea mucho más eficaz. Destaca el Afatinib, el cual se une covalentemente al residuo p.C797 del dominio TK cuando está activo. Esta unión se establece mediante la formación de un enlace entre la amida del compuesto y el grupo sulfuro de la cisteína (Solca *et al.*, 2012).

- 3ª generación: frente a la aparición de la mutación p.T790M en *EGFR*, se han desarrollado fármacos capaces de superar y aprovecharse de esta mutación. Recientemente aprobado se encuentra el compuesto AZD9291 (Cross *et al.*, 2014), un compuesto mono-anilino-pirimidínico. Esta molécula presenta una notable afinidad por las moléculas de *EGFR* con el cambio p.T790M, uniéndose irreversiblemente en el sitio de unión del ATP mediante un enlace covalente al residuo p.C797.

Debido a los tratamientos con ITK tanto de 1ª como de 2ª generación en pacientes con *EGFR* mutado, las células desarrollan resistencia al tratamiento mediante la aparición de una mutación en el codón 790, en concreto la p.T790M, que reduce drásticamente la afinidad de los fármacos por *EGFR* mutados, encontrándose en el 50% de los casos. La mutación produce un aumento de la afinidad por el ATP, reduciendo la que presentaba por los ITK. Otros mecanismos de resistencia incluyen la amplificación de *EGFR* (8%), mutaciones activantes en el gen *PIK3CA* (5%), amplificación del gen *MET* (5 -19%) y sobreexpresión del ligando GAS6 y su receptor *AXL* (25%). De esta forma, la célula tumoral es capaz de mantener la “adicción” a esta señalización (Garraway y Jänne, 2012; González de Castor *et al.*, 2015). Recientemente, se han reportado casos de resistencia a inhibidores de tercera generación a través de mutaciones que afectan a la unión covalente del inhibidor a la molécula (Yu *et al.*, 2015) o mediante alguno de los mecanismos ya citados, como la amplificación de *MET* (Planchard *et al.*, 2015)

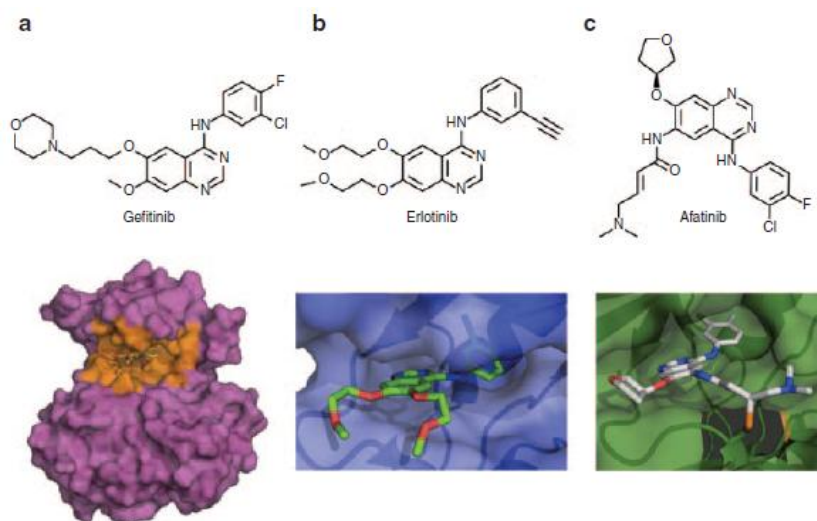


Figura 2: Ejemplos de estructura química de ITK y su unión a sus dianas. (A) - Estructura 2D del Gefitinib y unión por interacciones débiles al dominio tirosina cinasa de un EGFR con mutación activante (PDB code 3UG2). (B) - Estructura 2D del Erlotinib y estructura 3D de la interacción con un EGFR con mutaciones activantes (PDB code 4HJO). (C) - Estructura 2D del Afatinib y estructura 3D de su unión covalente a p.C797 (en color naranja) a EGFR con la mutación p.T790M (PDB code 4G5J). Tomada de Gonzalez de Castro *et al.* (2012)

Asimismo, los pacientes con CPNM portadores de traslocaciones en *ALK* poseen una medicina de precisión con ITK (Crizotinib). También se han descrito mecanismos de resistencia a



Crizotinib en las células con formas aberrantes de ALK, los cuales son muy parecidos a los de EGFR en cuanto a fundamento: la estrategia se basa en seguir manteniendo la actividad pese a la presencia del inhibidor. Así, ALK adquiere resistencia mediante amplificación génica, mutaciones en el dominio tirosina cinasa que reduzcan la afinidad de la molécula por el inhibidor o la aparición de nuevos oncogenes *driver* que permitan mantener el fenotipo tumoral (Doebele *et al.*, 2012; Garraway y Jänne, 2012).

## 1.5.Detección de biomarcadores

### 1.5.1. Detección de mutaciones en *EGFR*

Algunos de los métodos empleados en el diagnóstico y determinación de las principales mutaciones en el gen *EGFR* son (Lindeman *et al.*, 2013):

- Detección durante la PCR: la qPCR o la amplificación específica de alelo son técnicas que permiten detectar con una simple amplificación del gADN la presencia de mutaciones específicas.
- Detección post-PCR: la electroforesis capilar, la *Next Generation Sequencing* (NGS) o la pirosecuenciación son técnicas que permiten saber con precisión, a partir de una muestra previa y adecuadamente amplificada, cuál es la mutación que ha sucedido. Entre ellas variará la sensibilidad de cada método.
- Escaneo de mutaciones: el *High Resolution Melting* (HRM) o el *High Pressure Liquid Chromatography* (HPLC) en condiciones desnaturalizantes van a permitir ver si una muestra presenta o no mutación. Aunque estas técnicas solo informan de si hay mutación o no, permiten descartar las muestras no mutadas antes de procesarlas por otra técnica para identificar la mutación.
- Enriquecimiento del alelo mutante: la *co-amplification at lower denaturation temperature* (COLD) PCR o la *Peptide Nucleic Acid (PNA) Amplification* pueden ser técnicas útiles cuando la muestra carece de una población tumoral elevada. Mediante la amplificación favorable (no específica) de la variante mutada, se pretende igualar las proporciones de ADN mutado y ADN *wild-type* (wt).

### 1.5.2. Detección de reordenamientos en *ALK*

Para la detección de los transcritos de fusión del gen *ALK*, se emplean actualmente las siguientes técnicas (Tuononen *et al.*, 2013):

- Fluorescence *In Situ* Hybridization (FISH): mediante el empleo de parejas de sondas (sondas de rotura), se visualiza si ambas están lo suficiente separadas debido a la presencia de un reordenamiento. El par de sondas está diseñada para detectar reordenamientos concretos.
- Real-Time PCR: partiendo del ADNc de la muestra, se puede emplear cebadores y sondas que permitan detectar si existe amplificación o no de un reordenamiento

concreto. Es importante definir a partir de que ciclo umbral ( $C_t$ ) se considera que se ha producido el reordenamiento.

- Inmunohistoquímica: este ensayo se dirige hacia la proteína aberrante, empleando para ello una muestra fijada y un anticuerpo monoclonal dirigido específicamente contra ALK aberrantes.
- Secuenciación masiva o *Next-Generation Sequencing* (NGS): el uso de los cebadores adecuados permite la secuenciación de cualquier forma de ALK. La secuenciación puede ir dirigida tanto contra el ADN como contra los transcritos. Hay que tener en cuenta que en un único análisis podemos secuenciar también otros genes de interés.

## 1.6.Secuenciación masiva

El desarrollo de las tecnologías de secuenciación de ADN ha permitido un gran avance desde la secuenciación del primer genoma humano, en el cual fue necesaria la intervención de numerosos grupos de investigación y la inversión de un gran capital y de tiempo. En la actualidad, la secuenciación de un genoma humano requiere mucho menos tiempo y es mucho menos costosa (Bubnoff, 2008; Tucker *et al.*, 2009). En la última década, se ha demostrado que el cáncer puede ser debido al exceso de actividad de las proteínas derivadas de oncogenes o a la ausencia de los denominados TSG, desembocando en que la célula adquiera características tumorales. Es necesario, por tanto, identificar y distinguir las mutaciones que aparecen en los tumores, y, teniendo en cuenta la inestabilidad genómica característica de los tumores, cuáles de ellas dotan a los tumores de su fenotipo (mutaciones *drivers*) y cuáles no (mutaciones *passenger*) (Volsteing *et al.*, 2013). El desarrollo tecnológico de la secuenciación de última generación (NGS) nos permite obtener toda esta información tanto a nivel de *whole genome sequencing* (WGS), *whole exome sequencing* (WES) o incluso *targeted sequencing*.

En general, el fundamento de todas las NGS es la capacidad de llevar millones de reacciones de amplificación de cadena de forma paralela junto a un método de identificación del nucleótido introducido en la cadena (Bubnoff, 2008). Así, la pirosecuenciación introduce nucleótidos en cada ciclo, detectando luz en un pocillo solo si se incorpora a la cadena; la secuenciación mediante terminadores de Illumina detecta el fluoróforo del dNTP introducido en cada *spot* de cadenas sobre un chip (Tucker *et al.*, 2009); y la secuenciación por semiconductores introduce un dNTP en cada ciclo, detectando la incorporación del nucleótido a través de las variaciones de pH producidas en la polimerización (ThermoFisher, 2014).

Algunas ventajas presentadas por las NGS son que permiten generar gran cantidad de información bajo un mismo procedimiento; necesitan muy poca muestra para proceder, pudiendo algunas partir de 10ng de muestra en condiciones adecuadas para generar una secuencia consistente, y el tiempo empleado en la reacción es menor que la determinación de cada uno de los biomarcadores individualmente (Tucker *et al.*, 2009; Hovleson *et al.*, 2015).

Pero también vamos a encontrar algunas dificultades en el empleo de las NGS: protocolos y técnicas que requieren de personal con experiencia y preparados, incapacidad de distinguir las mutaciones *driver* de las *passenger* ante la presencia de nuevas mutaciones, entender cómo la

mutación afecta a la actividad de la proteína o la pérdida de datos durante el procesamiento debido a errores informáticos (Tucker *et al.*, 2009; Volsteing *et al.*, 2013; Raphael *et al.*, 2014).

### 1.7. Secuenciación dirigida en CPNM

Cuando se secuencia el genoma de un paciente con una determinada enfermedad, existen estudios previos de qué alteraciones están relacionadas con la enfermedad del paciente. En un caso como este, no es necesario secuenciar los más de 25.000 genes humanos, siendo más útil enfocar la secuenciación en los genes ya previamente correlacionados con la patogénesis. Estas propuestas de secuenciación, denominadas en general como *Targeted Sequencing*, pretenden extraer únicamente la información necesaria o solicitada del genoma a secuenciar.

En la actualidad, gracias a la labor de consorcios internacionales como el *The Cancer Genome Atlas* (TCGA) o a bases de datos como *Oncomine* (Rhodes *et al.*, 2007), se conocen todo tipo de alteraciones relacionadas con determinados tipos de cáncer. En función de las necesidades de información a obtener, la muestra requerida y las capacidades de los equipos, entre otros factores, en oncología se plantean varios tipos de ensayos (Hovelson *et al.*, 2015):

- Secuenciación del genoma, exoma o transcriptoma
- Paneles basados en hibridación
- Paneles basados en amplificación mediante PCR

Uno de los kits disponibles en el mercado, para la plataforma PGM de Ion Torrent (ThermoFisher, USA), es el de *Oncomine Solid Tumour DNA kit* (ThermoFisher, USA). Este kit permite la secuenciación de una serie de 22 genes implicados tanto en CP como cáncer de colon. Alfabéticamente, estos genes son: *AKT1*, *ALK*, *BRAF*, *CTNNB1*, *DDR2*, *EGFR*, *ERBB2*, *ERBB4*, *FBXW7*, *FGFR1*, *FGFR2*, *FGFR3*, *KRAS*, *MAP2K1*, *MET*, *NOTCH1*, *NRAS*, *PIK3CA*, *PTEN*, *SMAD4*, *STK11*, *TP53*. Estos genes son testados a través de 92 amplicones que permiten un *Targeted sequencing* a partir de 10 ng de gADN del paciente.

Para la secuenciación del ARN, existe el panel *Oncomine Solid Tumour Fusion Transcript* (ThermoFisher, USA), el cual permite, a través de 85 amplicones, determinar la presencia de diferentes transcritos de fusión de los genes *ALK*, *RET*, *ROS1* y *NTRK1* a partir de 10 ng de ARN extraído de las células del paciente.

El CPNM es una enfermedad heterogénea, donde coexisten tumores formados a través de distintas vías de carcinogénesis y, como se ha comentado, el tratamiento personalizado del CPNM se dirige a través del estatus mutacional del gen *EGFR* o reordenamientos de *ALK*. La presencia de estas alteraciones identifica a pacientes con tumores de pulmón sensibles al tratamiento con ITK. Sin embargo, aproximadamente un 20-25% de los pacientes así seleccionados no responden a la terapia. La respuesta a esta divergencia debe buscarse en la heterogeneidad intratumoral, en la diversidad entre el tumor primario y sus metástasis, y en la sensibilidad de los métodos empleados para identificar las mutaciones. El cáncer es un proceso dinámico y es en este contexto en el que se deben considerar los conceptos de evolución clonal y heterogeneidad intratumoral, así como valorar la repercusión clínica que conllevan en una aproximación personalizada del tratamiento de esta enfermedad, creemos que

tecnologías como la secuenciación masiva pueden ofrecer información relevante para adecuar el tratamiento del paciente.

## 2. Objetivos

Este Trabajo Final de Grado pretende comparar los diferentes métodos actuales empleados en CPNM en la labor asistencial del Laboratorio de Biología Molecular del Hospital Universitario y Politécnico La Fe (HUyP) y discutir el alcance y las limitaciones de cada uno de ellos. Para la consecución de este fin, se proponen los siguientes objetivos:

- Determinación de las alteraciones moleculares en el gen *EGFR* en pacientes con CPNM mediante la utilización de tres métodos: Secuenciación directa, qPCR mediante Cobas Analyzer y *targeted sequencing* mediante Personal Genome Machine (PGM; ThermoFisher, USA).
- Comparación de la concordancia de los resultados obtenidos a través de los diferentes métodos empleados.

## 3. Material y métodos

### 3.1. Muestras

Las muestras empleadas para este estudio retrospectivo corresponden a biopsias sólidas de tumores localizados en pulmón o de posibles metástasis de los mismos de pacientes atendidos por el Servicio de Oncología Médica del HUyP La Fe entre los años 2015-2016. Todas estas muestras han sido autorizadas para su empleo en este proyecto a través del consentimiento informado y firmado por los pacientes, mediante el cual ceden el excedente de muestra tras finalizar el análisis para su uso en investigación, autorizado por el Comité de Ética del Hospital La Fe.

Tras la toma de la biopsia, el servicio de Anatomía Patológica se ha encargado de la conservación de la muestra, mediante fijación en formol e inclusión en bloques de parafina (*Formalin-fixed, Paraffin-embedded* (FFPE)) y han realizado los cortes de 5-10 µm necesarios para los estudios moleculares en ADN y ARN. Los requisitos de representación tumoral (>50%) y celularidad tumoral total (> 150-300 células) han sido establecidos por el patólogo responsable en todos los casos.

El Laboratorio de Biología Molecular del Servicio de Análisis Clínicos del HUyP La Fe ha recibido un total de 40 muestras para este estudio. Como algunas muestras correspondían a segundas biopsias o metástasis (Anexo I), el total de pacientes es de 36. Aquellas muestras pertenecientes a un mismo paciente, han sido identificadas por números con misma parte entera pero diferente decimal referenciando orden de llegada.

## 3.2.Extracción de ácidos nucleicos

### 3.2.1. Extracción de ADN

El ADN de las muestras es aislado mediante el *GeneRead DNA FFPE kit* (QUIAGEN, Alemania). En el laboratorio se reciben cortes o “punchs” de la muestra parafinada. A cada muestra a extraer se le añade un volumen superior a 160 µl de Solución de Desparafinado que cubra la muestra. Tras vortear, se incuban a 56°C unos 3min. Antes de añadir 100 µl a cada muestra de una mezcla compuesta de 55 µl *RNase-Free wáter*, 25 µl Buffer FTB y 20 µl de proteinasa K, se espera a que la muestra se atempere. Una vez añadido y vorteadado, se deja calentar toda la noche a 56°C. A primera hora de la jornada, se traspasa a un bloque ya caliente a 90°C durante una hora. Sin exceder el tiempo, se retira y se da un golpe de centrifuga a la muestra. Se observará una fase superior azul con la parafina y una transparente inferior. Esta segunda fase se transferirá un nuevo tubo de 1,5 ml. Se añaden 115 µl de *RNase-Free water* y 35 µl de Uracil-N-Glycolasa, un enzima que revierte las modificaciones de las citosinas causadas por el proceso de inclusión en parafina (Figura 3). La mezcla se incuba a 50°C durante una hora.

A continuación, se añaden 2 µl de *RNase A* y se deja incubar 2 min a Tª ambiente. Al pasar el tiempo, se añaden 250 µl de Buffer AL y 250 µl de etanol puro. Tras vortear y dar un golpe de centrifuga, se pasa un máximo volumen de 700 µl por una columna de purificación. Si hubiera un exceso, debería emplearse una segunda columna. Se centrifuga un minuto a máxima velocidad (13000 rpm apróx.) y se descarta el fluido. A continuación se lava la muestra tres veces, añadiendo una solución y centrifugando a máxima velocidad durante un minuto, descartando después el fluido filtrado. Primero se lava con 500 µl de Buffer AW 1 y luego con 500 µl de Buffer AW 2. La columna se transfiere a un tubo nuevo para eliminar posibles residuos mediante centrifugación. Finalmente, se pasa la columna a un nuevo tubo de 1,5 ml, y se le añaden 30 µl de Tris-EDTA (pH 7,5, Tris 10 mM, EDTA 1 mM). Tras una breve incubación, se centrifuga un minuto a máxima velocidad y se descarta la columna. Se realizarán dos eluciones en dos tubos diferentes.

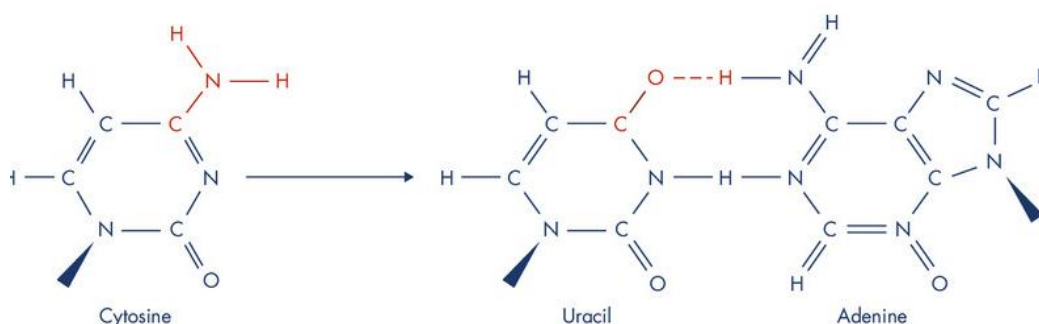


Figura 3: Deaminación oxidativa de la citosina en muestras FFPE. El ambiente en el que se encuentra una muestra FFPE da lugar a esta modificación química. Este cambio producirá un mismatch en la cadena, que en el momento de una PCR dará lugar a la fijación de la mutación C>T. Tomada de [www.quiagen.com](http://www.quiagen.com).

Con el fin de conservar el ADN adecuadamente, el volumen se traspasa a tubos de 200 µl debido a su comodidad para el almacenamiento. En caso de uso inmediato, las muestras se conservan a 4°C no más de una semana. Para tiempos más prolongados se mantienen en congeladores a -80°C.

### 3.2.2. Extracción de ARN

El ARN se extrae a partir de cortes procedentes del mismo bloque de parafina que el ADN correspondiente a la misma muestra. Para su extracción, se emplea el *RecoverAll™ Total Nucleic Acid Isolation Kit* (ThermoFisher, USA). El proceso se inicia con la adición de un mínimo de 100 µl de Solución de desparafinado que recubran completamente la muestra y calentándolo a 50°C durante 3 min. Se añaden 100 µl de *Digestion Buffer* y se vortea todo el contenido del tubo. En la fase transparente que se forme, se añadirán los 4 µl de *Protease* y se mezcla por pipeteo. Se incuba en bloque seco 3 h a 50°C y 15 min a 80°C. Durante el proceso, se prepara un mix que contendrá un volumen de 395 µl por tubo, componiéndose de 120 µl de *Isolation additive* y 275 µl de etanol puro. Se añade el volumen a cada tubo, y se mezcla por pipeteo. La mezcla se pasa por una columna de purificación, mediante centrifugación 30 s a 10000 rpm. A continuación, se realiza un primer lavado añadiendo 700 µl de *Wash 1*, centrifugando 30 s a 10000 rpm y descartando el filtrado. El lavado se repite una vez con 500 µl de *Wash 2/3* y una última vez sin añadir solución de lavado para eliminar los restos que hayan quedado en la columna.

En este paso de la extracción, la columna retiene ADN y ARN. Para conservar el ARN, se añade a cada columna 60 µl de una solución compuesta por 6 µl de *10X DNase Buffer*, 4 µl de DNasa y 50 µl de agua libre de RNasas. Se deja incubar en la membrana 30 min a Tª ambiente. Para eliminar la DNasa, se repiten los lavados con 700 µl de *Wash 1*, dos veces con 500 µl de *Wash 2/3* y un último lavado sin solución para descartar los desechos restantes. Para eluir la muestra, se transfiere la columna a un nuevo tubo, se adicionan 30 µl de agua libre de nucleasas y se incuba 5-10 min antes de centrifugar un minuto a máxima velocidad.

El ARN se conserva a -80°C mientras no vaya a ser empleado. En caso de que se vayan a hacer librerías de ARN en el mismo día, se conserva a 4°C hasta que se inicia la reacción.

### 3.3. Cuantificación de ácidos nucleicos

Para conocer la cantidad y calidad de ADN, las muestras son cuantificadas mediante *Qubit 3.0 fluorometer* (ThermoFisher, USA). Debido a la concentración con las que se obtienen las muestras tras la extracción, se emplea el *Qubit dsDNA HS Assay kit* (ThermoFisher, USA). El primer paso del ensayo será preparar solución de trabajo para dos estándares y dos cuantificaciones por muestra a partir de 1 µl de *Qubit dsDNA HS reagent* diluido en 199 µl de *dsDNA HS buffer*. Una vez preparado, se guarda en oscuridad. Los estándares se preparan añadiendo 10 µl de cada estándar a 190 µl de solución de trabajo. Nada más iniciar el *Qubit 3.0 fluorometer* e indicar que queremos cuantificar dsADN, en primer lugar se requiere calibrar el aparato midiendo la concentración en ambas soluciones conforme lo solicite. Para preparar las

muestras, se diluye 1 µl de esta en 199 µl de solución de trabajo. Rápidamente, se coloca la muestra en el aparato y se determina la concentración del ADN extraído.

Para cuantificar ARN, se utiliza el *Qubit RNA HS Assay kit* (ThermoFisher, USA), siguiendo el mismo protocolo pero con los reactivos correspondientes del kit.

### 3.4.Detección de mutación mediante PCR en tiempo real (plataforma Cobas®)

La determinación de la presencia de mutaciones en los exones 18 a 21 del gen *EGFR* se realizará mediante el kit *Cobas EGFR Mutation test v2* (Roche, Suiza). A partir del ADN extraído, se preparan diluciones de la muestra para obtener una concentración de 2 ng/µl en un volumen mínimo de 90 µl. Las diluciones se realizan con el *Speciment Diluent* del kit, en base a una planilla que realiza los cálculos de los volúmenes necesarios. En caso de que las muestras excedan los 70 ng/µl, para no malgastar *Speciment Diluent*, se trabaja como si no excediera el límite y se emplean 5µl de muestra y 170 µl de diluyente. Una vez preparadas, se enciende el *Cobas 480z Analyzer* con la finalidad de inicializar el aparato.

Se preparan 3 mixes de las tres soluciones del kit, mezclándose 20 µl del MMX1, MMX2 o MMX3 con 5 µl de MGAC. Para cada mix se preparan 75 µl más 25 µl por muestra. A continuación se carga la placa adecuadamente (Figura 4). Una vez cargada, se sella la placa con un adhesivo y se carga en el *Cobas 480z Analyzer*. Por último, se termina de configurar el ensayo indicando el kit empleado, las muestras a analizar y la identificación de las mismas. El instrumento reportará tras hora y media de reacción las mutaciones detectadas, aunque sin indicar la posición correspondiente en el ADNc. En la tabla 1 se indican las mutaciones detectadas con el kit.

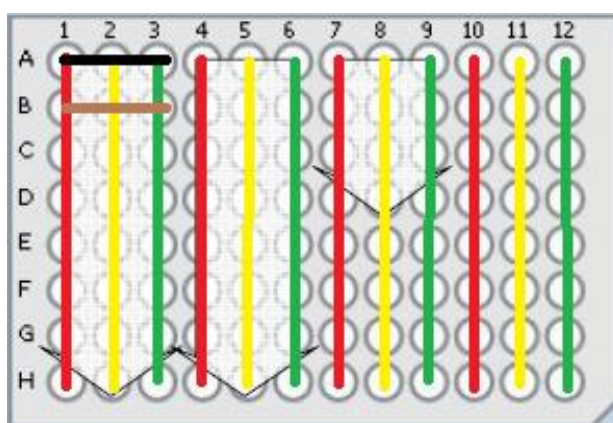


Figura 4: Esquema de la carga de una placa de 96 pocillos para el *Cobas 480z Analyzer*. En orden, se llenarán dos pocillos más el equivalente al número de muestras con 25 µl de MMX1 (rojo), MMX2 (amarillo) y MMX3 (verde). Los pocillos A1, A2 y A3 (negro) se llenan con el control Mutado, los B1, B2 y B3 (marrón) con *Speciment Diluent*, y los siguientes grupos de tres (C1, C2 y C3; D1, D2 y D3...) con 25 µl de la muestra preparada

Tabla 1: Mutaciones de EGFR detectadas por *Cobas EGFR Mutation test v2*

| Exón    | Mutación en EGFR | Variación en el gen <i>EGFR</i> | COSMIC ID |
|---------|------------------|---------------------------------|-----------|
| Exón 18 | G719X            | 2156G>C                         | 6239      |
|         |                  | 2155G>A                         | 6252      |
|         |                  | 2155G>T                         | 6253      |
| Exón 19 | Ex19Del          | 2240_2251del12                  | 6210      |
|         |                  | 2239_2247del9                   | 6218      |
|         |                  | 2238_2255del18                  | 6220      |
|         |                  | 2235_2249del15                  | 6223      |
|         |                  | 2236_2250del15                  | 6225      |
|         |                  | 2239_2253del15                  | 6254      |
|         |                  | 2239_2256del18                  | 6255      |
|         |                  | 2237_2254del18                  | 12367     |
|         |                  | 2240_2254del15                  | 12369     |
|         |                  | 2240_2257del18                  | 12370     |
|         |                  | 2239_2248TTAAGAGAAG>C           | 12382     |
|         |                  | 2239_2251>C                     | 12383     |
|         |                  | 2237_2255>T                     | 12384     |
|         |                  | 2235_2255>AAT                   | 12385     |
|         |                  | 2237_2252>T                     | 12386     |
|         |                  | 2239_2258>CA                    | 12387     |
|         |                  | 2239_2256>CAA                   | 12403     |
|         |                  | 2237_2253>TTGCT                 | 12416     |
|         |                  | 2238_2252>GCA                   | 12419     |
|         |                  | 2238_2248>GC                    | 12422     |
|         |                  | 2237_2251del15                  | 12678     |
|         |                  | 2236_2253del18                  | 12728     |
|         |                  | 2235_2248>AATTC                 | 13550     |
|         |                  | 2235_2252>AAT                   | 13551     |
|         |                  | 2235_2251>AATTC                 | 13552     |
|         |                  | 2253_2276del24                  | 13556     |
|         |                  | 2237_2257>TCT                   | 18427     |
|         |                  | 2238_2252del15                  | 23571     |
|         |                  | 2233_2247del15                  | 26038     |
| Exón 20 | S768I            | 2303G>T                         | 6241      |
|         | T790M            | 2369C>T                         | 6240      |
|         | Ex20Ins          | 2307_2308ins9GCCAGCGTG          | 12376     |
|         |                  | 2319_2320insCAC                 | 12377     |
|         |                  | 2310_2311insGGT                 | 12378     |
|         |                  | 2311_2312ins9GCGTGGACA          | 13428     |
| Exón 21 | L858R            | 2309_2310AC>CCAGCGTGGAT         | 13558     |
|         |                  | 2573T>G                         | 6224      |
|         | L861Q            | 2573_2574TG>GT                  | 12429     |
|         |                  | 2582T>A                         | 6213      |



### 3.5.Detección de mutación mediante PCR convencional y Secuenciación directa

En el presente proyecto, las muestras que han dado positivo para una mutación en Cobas han sido confirmadas mediante secuenciación directa. Para ello se realiza la amplificación del exón que presenta la mutación. La PCR se realiza mediante el kit *Type-it mutation detection* (Roche, USA). Cada tubo de reacción se compondrá de 12,5 µl de *Type-it multiplex PCR master mix X2*, 2,5 µl de *Q-solution X5*, 0,5 µl del cebador *forward* del exón correspondiente y 0,5 µl del cebador *reverse*, ambos a una concentración 10 µM (Tabla 2), y un volumen de 9 µl compuesto por el volumen necesario de muestra para alcanzar 100ng de muestra y *Rnase-Free water* hasta completar el volumen. Por cada grupo de muestras analizadas simultáneamente (2 a 4), se preparara un blanco en el que los 9µl serán agua para comprobar que no hay contaminación y un control negativo previamente testado en el que no se detecte ninguna mutación. La PCR consta de los siguientes pasos: Desnaturalización inicial a 95°C durante 5 minutos; 40 ciclos, cada uno de ellos compuesto por Desnaturalización a 95°C 45s, Hibridación a 57°C 45s y Elongación a 72°C 1min; Extensión final a 72°C durante 10 min.

Tabla 2: Cebadores empleados para los distintos exones del gen EGFR.

| Exón | Sentido | Secuencia                  | Tª (°C) | Amplicón (pb) |
|------|---------|----------------------------|---------|---------------|
| 18   | Forward | CAAGTGCCGTGTCCTGGCACCCAAGC | 57      | 372           |
|      | Reverse | CCAAACACTCAGTGAAACAAAGAG   | 57      |               |
| 19   | Forward | GCAATATCAGCCTTAGGTGCGGCTC  | 57      | 366           |
|      | Reverse | CATAGAAAGTGAACATTTAGGATGTG | 57      |               |
| 20   | Forward | TTTGAAACTCAAGATCGCATT      | 57      | 368           |
|      | Reverse | CATATCCCCATGGCAAACCTC      | 57      |               |
| 21   | Forward | CAGCCATAAGTCCTCGACGTGG     | 57      | 372           |
|      | Reverse | CATCCTCCCCTGCATGTGTAAAC    | 57      |               |

Una vez finalizada la PCR, se comprueba que la amplificación haya sido correcta mediante electroforesis en gel de agarosa. Para ello, se diluye la muestra con agua en proporción 3:17. También se prepara una dilución 1:9 del marcador de pesos moleculares *TrackIt 50 bp Ladder DNA* (ThermoFisher, USA). Con las diluciones listas, se coloca el gel *E-gel Ex 2% agarose* (ThermoFisher, USA) en el *E-Gel Agarose System* (ThermoFisher, USA). El gel, consta de 11 pocillos, uno para el marcador y el resto para muestras. Para cargar el gel, la dilución del marcador se pipetea directamente en el pocillo de la derecha. En cambio, antes de ser cargadas, las muestras se mezclan con 1µl de tampón de carga. Una vez está listo el gel, se selecciona en el aparato el gel empleado y se deja correr durante 12 min. Una vez finalizado, se toma una imagen del gel.

Una vez comprobado que la amplificación ha sido correcta, se prepara un nuevo tubo por cada muestra con 2 µl del enzima ExoSAP y 8 µl del producto de PCR, incubado durante 15 min a 37°C y 15 min a 80°C. Tras este paso de eliminación de exceso de cebadores no consumidos en la PCR, se procede con la reacción de secuencia. Por cada tubo de muestra, se preparan dos tubos, uno para la secuencia forward y otro para la secuencia reverse. Cada tubo de la reacción

de secuencia se compone de 3,5 µl de H<sub>2</sub>O libre de RNasa, 1 µl de *Terminator Ready reaction mix*, 0,5 µl del cebador correspondiente y 5 µl de la muestra purificada por ExoSAP. El programa de PCR empleado consta de 25 ciclos de tres pasos: Desnaturalización a 96°C unos 10s, Hibridación a 50°C por 5s y Extensión de la cadena 4 min a 60°C.

Por último, antes de cargar en placa, se purifica el producto de la PCR a través de las columnas de purificación *Quiamp Minielute Column* (QUIAGEN, Alemania). Debido a que las columnas se conservan hidratadas, en una primera centrifuga 2 min a 3000 rpm se elimina el agua. La columna se traslada a un tubo de recogida, y ya se cargan los 10 µl de muestra. Tras una centrifugación en las mismas condiciones, se descarta la columna. La muestra es mezclada con 15 µl de formamida y se carga en placa de 96 pocillos apta para la secuenciación directa en el aparato *Genetic Analyzer 3100* (Applied Biosystem, USA). La placa cargada se centrifuga 1 min a 1500 rpm y se pone a desnaturalizar por calentamiento a 95°C durante 5 min. Una vez cargado en el secuenciador, se establece un programa que emplea el polímero POP7 y un tiempo de lectura de 30 min.

Finalmente, el archivo generado tras la secuenciación es procesado empleando el software *Sequencing Analysis 5.2* (Applied Biosystem, USA) y se analiza manualmente la secuencia. Debido a la baja sensibilidad comentada anteriormente, es necesario comprobar puntos clave de la secuencia en busca de picos que no hayan sido identificados por el secuenciador.

### **3.6.Detección de mutación mediante secuenciación masiva (plataforma PGM®)**

La secuenciación mediante semiconductores del *Personal Genome Machine* (PGM; ThermoFisher, USA) emplea unas secuencias *Barcode* que permiten identificar a las muestras con la finalidad de multiplexar el ensayo. Por ello, al disponer de 16 Barcodes diferentes, será posible secuenciar el ADN y ARN de hasta 8 pacientes en un único chip 318 v2.

#### **3.6.1. Genoteca de ADN**

Para preparar cada genoteca de ADN cada muestra es amplificada empleando el pool de cebadores y reactivos del *Oncomine Solid Tumor DNA kit* (ThermoFisher, USA). Este kit permite la detección de mutaciones en 22 genes mediante la generación de 96 amplicones. Para ello, se prepara en una placa de 96 pocillos 4 µl de *LIB HiFi Mix*, 4 µl de cebadores del panel 5X y un volumen total de 12 µl con 10 ng de muestra y el resto del volumen agua libre de nucleasas. Una vez añadido, se sella y se agita la placa. La reacción de PCR se lleva a cabo en el termociclador Verity Dx (ThermoFisher) empleando el siguiente programa: Desnaturalización inicial a 99°C durante 2 min, 30 ciclos de Desnaturalización a 99°C por 15 s y *Annealing* y Extensión durante 4 min a 60°C, conservación a 10°C. Finalizada la PCR, se añaden 2 µl de *LIB FuPa* para eliminar el exceso de cebadores. Tras homogeneizar, se introduce la placa en el Verity Dx donde se incubará 10 min a 50°C, otros 10 min a 55°C y 20 más a 60°C.

A continuación, se ligan los adaptadores y los *barcodes*. Para ello, se añade a cada tubo 6 µl de *LIB Switch* a cada muestra y 2 µl de barcodes, añadiéndose uno diferente a cada muestra. Para

dar comienzo a la reacción, se añaden 2 µl de *DNA ligase*, se sella la placa y se agita, y se coloca en el Verity Dx, con un programa de 22°C durante 20 min y 72°C durante 10 min.

Se aprovecha el tiempo de espera para preparar 330 µl de Etanol al 70% por cada muestra (2,640 ml para las 8 genotecas). Finalizada la ligación, se añaden 45 µl de *Agencourt AMPure XP* (Beckman Coulter, USA) a cada muestra y se mezcla por pipeteo, dejándolo incubar 5 min a Tª ambiente. La placa se coloca sobre un soporte imantado durante tres minutos y se retira el sobrenadante, con cuidado de no arrastrar el pellet, donde estarán las esferas con la muestra. Se lava dos veces con el etanol al 70%: se añaden 150 µl de etanol, se mezcla por pipeteo, se deja en el imán hasta que se forme el pellet y se descarta el sobrenadante. Por último, para eliminar el etanol restante, se deja evaporar a Tª ambiente.

El proceso de secuenciación requiere que las genotecas estén inicialmente a una concentración 100pM. Para ello, se procede a *Equalizar* las 8 muestras de ADN: esto es, llevarlas a dicha concentración. Para ello, se preparan 52 µl por muestra de un mix que contiene 2 µl de *LIB primers*, 10 µl de *LIB HiFi Mix* y 40 µl de agua libre de nucleasas. Tras la evaporación del etanol, se añade el volumen a cada genoteca y se mezcla por pipeteo. Tras sellar y centrifugar a 100 rcf por 30 s, se aplica el siguiente programa en el Verity Dx: 2 min a 98°C, 7 ciclos de Desnaturalización a 98°C por 15 s, y *Annealing* y extensión un minuto a 60°C, finalizando a 10°C.

Para purificar la muestra, se lavan 3 µl de *LIB beads* con 6 µl de *LIB Wash Solution* por muestra. Tras homogeneizar y centrifugar brevemente, se deja en el imán un minuto y se descarta el sobrenadante. El pellet se resuspende en 6 µl de *LIB Wash Solution*. Tras incubar 5 min, a la muestra se le añade 6 µl de *LIB beads* lavadas y dejando incubar 5 min. La placa se coloca en el imán durante 3 min, se descarta el sobrenadante y añaden 150 µl de *LIB Wash Solution*, que tras dejar que se forme el pellet, se descarta. Tras repetir el lavado, el pellet se resuspende en 100 µl de *LIB Elution Solution*. Tras homogeneizar, se calienta 5 min a 35°C. Se deja imantado durante 3 min, y el sobrenadante se traslada a un tubo en el que almacenar la librería. Para comprobar la concentración, se puede cuantificar por Qubit.

### 3.6.2. Genoteca de ARN

La producción de la genoteca de ARN es parecida a la de ADN, dándose un paso de síntesis de ADNc previo empleando el *SuperScript VILO cDNA síntesis kit* (ThermoFisher, USA). En este paso, se añade a cada pocillo de una placa de 96 un mix por muestra compuesto por 2 µl de *5X RT reaction Mix*, 1 µl de *10X enzyme Mix* y 7 µl que contengan un volumen de muestra con 10 ng y agua libre de nucleasas hasta completarlo. La placa sellada, vorteadada y centrifugada se introduce en el Verity Dx con un programa de 42°C 30 min, 85°C 5 min y finalizando a 10°C. La amplificación, limpieza de cebadores y ligación de adaptadores y *barcodes* sigue los mismos pasos que en la preparación de librerías de ADN. Hay dos aspectos a tener en cuenta:

- A los 10 µl de ADNc ya existentes en el pocillo, se añaden 2 µl de agua libre de nucleasas, 4 µl de *LIB HiFi Mix* y 4 µl del *pool* de cebadores 5X del *Oncomine Solid Tumour Fusion Transcript kit* (ThermoFisher, USA).

- Los *barcodes* empleados deben ser diferentes a los empleados con ADN, y únicos para cada muestra.

El procedimiento de lavado con Etanol al 70% también es igual al empleado con ADN. Tras la evaporación del etanol, se añaden 50 µl de *LIB Elution Mix*. Tras sellar, homogeneizar y aplicar un campo magnético 2 min, se retiran 5 µl de sobrenadante y se mezclan con 495 µl de agua libre de RNasas en tubos *Lo-bind* de 1,5 ml.

Para este tipo de librerías no es recomendable seguir el protocolo de ecualización para llevar la genoteca a 100pM, es su lugar se realiza la cuantificación de la genoteca por qPCR infiriendo la concentración de la misma a partir de una curva patrón. Para tal procedimiento, se empleará el *Ion Library Quantification Kit* (ThermoFisher, USA). El protocolo comienza preparando un volumen de 200 µl con una dilución 1:100 de la librería mediante la mezcla de 2 µl de muestra con 198 µl de agua libre de nucleasas y tres diluciones decimales seriadas del estándar, obteniéndose uno a 6,8 pM, otro a 0,68 pM y un tercero a 0,068 pM. Por cada librería y por los tres estándar se prepara una solución de 22µl compuesta por 20 µl de *2X TaqMan Master Mix* y 2 µl de *20X TaqMan Assay*. Por cada muestra y estándar, se prepararán dos pocillos que se llenarán con 11 µl de la solución y 9 µl de la muestra diluida o el estándar. Una vez, cargada la placa, se coloca en el *Applied Biosystems 7500 Real-Time PCR System* (ThermoFisher, USA). El programa de PCR a emplearse compone de un paso inicial de Desnaturalización a 95°C durante 2 min y 40 ciclos de 15 s a 95°C y 1 min a 60°C. Tras conocer la concentración de cada librería se calcula el factor de dilución correspondiente para llevar cada genoteca a 100 pM con la adición de la cantidad de *Low-TE* necesaria.

### 3.6.3. Combinar librerías

La utilización de *barcodes* permitirá identificar la genoteca de origen cuando las muestras sean conjuntamente secuenciadas. Esto permite que las 16 genotecas preparadas se puedan cargar en un mismo chip. Para generar las librerías combinadas, se juntan 10 µl de cada genoteca de ADN en un tubo y 10 µl de las de ARN en otro tubo. En un tercer tubo, se añaden 16 µl de la mezcla de genotecas de ADN y 4 µl de las de ARN.

### 3.6.4. PCR en emulsión

La finalidad de esta etapa es generar las microesferas con la secuencia que va a ser leída. Todos los reactivos de este paso pertenecen al *Ion PGM Hi-Q OT2 kit* (ThermoFisher, USA). Antes de comenzar, hay que encender el *Ion OneTouch 2 Instrument* (ThermoFisher, USA) y prepararlo según las instrucciones del fabricante. En este paso, en el aparato se introducen los dos tubos de colección.

La librería combinada se diluye mediante adición de 23 µl de agua libre de nucleasas a 2 µl de la librería 100 pM. A este volumen se le añaden 800 µl de *Ion PGM Hi-Q Reagent Mix*, 25 µl de agua libre de nucleasa, 50 µl de *Ion PGM Hi-Q Enzyme Mix* y 900 µl de *Ion PGM Hi-Q ISP*.

Con el mix ya listo, se procede a cargar el aparato y a preparar el *Ion OneTouch Reaction Filter*, donde se cargarán los reactivos necesarios, entre los que se encuentran los *Ion Sphere Particle* (ISP) y la muestra. En la apertura para la muestra, se le añade todo el mix preparado, más 1,7 ml de *Ion OneTouch Reaction Oil*. A continuación, se coloca el filtro en el aparato y este se hace funcionar bajo el programa PGM: *Ion Torrent PGM Hi-Q OT2 kit 200*. A partir de aquí, el proceso dura unas 5 horas hasta un spin final de 10 min. Transcurrido este tiempo, la muestra se habrá depositado en ambos tubos. Para recuperarla, se elimina en ambos tubos todo el sobrenadante salvo 100 µl en cada uno. Este volumen se utiliza para resuspender el pellet. Después se le añade a cada tubo 500 µl de *Ion OneTouch Wash Solution* y se mezcla por pipeteo. El volumen de los dos tubos se transfiere a un Eppendorf *Lo-bind* de 1,5 ml, se centrifuga 2,5 min a 15000 rcf y se descarta todo el pellet salvo 100 µl que se utilizan para resuspender la muestra.

### 3.6.5. Enriquecimiento de la muestra en esferas con PCR clonal

Este paso pretende reducir al mínimo el número de esferas sin fragmento que entran en el chip, ya que lo único que hacen es competir por los pocillos con las muestras cargadas. Este proceso tiene lugar en el *Ion OneTouch ES* (ThermoFisher, USA) en una tira de compuesta por 8 pocillos que se cargan de la siguiente manera:

- Pocillo 1: Muestra resuspendida en los 100 µl de sobrenadante.
- Pocillo 2: 130 µl de *Dynabeads MyOne Streptavidin C1 Beads*. Tras sacarlas del congelador, las esferas se resuspenden y centrifugan. Estas se transfieren a un nuevo tubo *Lo-bind*. Mediante un imán, se genera un pellet y se descarta el sobrenadante. Se añaden 130 µl de *MyOne Wash Solution*, se mezcla y se transfiere al pocillo.
- Pocillo 3, 4 y 5: 300 µl *Ion OneTouch Wash Solution* en cada uno.
- Pocillo 6 y 8: vacío.
- Pocillo 7: 300 µl de solución *Melt-off*. La solución se prepara con 280 µl de *Tween solution* y 40 µl de NaOH 1 M, produciéndose 20 µl de exceso.

Una vez cargada la tira, se coloca una nueva punta al aparato y un tubo de PCR de 0,2 ml con 10 µl de *Neutralization solution*. Antes de inicializar, es recomendable resuspender las esferas. El proceso de enriquecimiento dura 35 minutos, transcurrido este tiempo se retira el tubo con la librería enriquecida y se descarta la tira. Este tubo se puede almacenar hasta 3 días a una temperatura de 2-8°C.

### 3.6.6. Carga del Chip y Secuenciación

El primer paso antes de cargar el Chip que contiene los semiconductores y los pocillos es programar el *run* a través del *Torrent Browser* (ThermoFisher, USA). Es importante identificar cada muestra con su *barcode* correspondiente, además de indicar que las genotecas de ADN se compararán con el genoma de referencia y las ARN con un archivo que contiene los transcritos de fusión que detecta el kit. A continuación, se inicializa el PGM, siguiendo las instrucciones

aportadas por el aparato. En este paso, se cargarán las distintas soluciones de lavado y reactivos para ajustar el PH de la reacción y las 4 soluciones de nucleótidos en el aparato.

Antes de comenzar la preparación de las ISP enriquecidas, se les añade 5 µl de *Test Fragment*, utilizados por la PGM para determinar la calidad de la carrera. Las ISP se homogeneizarán y se centrifugarán 2 min a 15500 rcf. Se descarta todo el sobrenadante salvo 15 µl (aproximadamente), a los que se añaden 12 µl de cebadores y el volumen suficiente de *Annealing buffer* (en caso de haber dejado menos de 15 µl). La muestra se calentará 2 min a 95°C y 2 min a 37°C. Por otro lado, se coloca en el PGM un chip 318 v2 para el *Chip Check*, un proceso que comprueba el buen estado del chip. Mientras este tiene lugar, se añaden 3 µl de *Ion PGM Sequencing 200 v2 Polymerase* a los 27 µl de ISPs enriquecidas.

Cuando trascurran 5 min de la adición de la polimerasa y el *Chip Check* haya finalizado, se vacía el líquido que este contenido en el chip. Tras centrifugarlo, se le añaden los 30 µl de ISPs por una ranura del chip, evitando la entrada de aire. El exceso de volumen se retira y se centrifuga 30 s. Para mezclar la solución, se extraen 20 µl del chip, se vuelven a introducir y se centrifuga 30 s. Una vez se repita este proceso se descarta el exceso. Por último, el chip, ya preparado, se carga en el PGM y se inicia el proceso planificado.

### 3.7. Análisis bioinformático

La secuenciación genera un archivo UBAM con todas las lecturas obtenidas en los pocillos. Este archivo es subido al programa *Ion Reporter* (ThermoFisher, USA) para su análisis y mapeo, con el fin de construir la secuencia de los amplicones. En el Ion Reporter se aplica el programa de análisis correspondiente al kit empleado que determinará las variantes encontradas en la muestra. Para considerar que la carrera ha sido valida, deben darse un mínimo de 100000 lecturas en genotecas de ADN y 20000 en las de ARN. Además, la cobertura mínima de una mutación debe ser de 20X para considerar que no es ruido de fondo. En caso de considerarlo necesario o de querer profundizar en las lecturas de una genoteca, se procede al visualizado del archivo BAM mediante Integrative Genomics Viewer (IGV; Broad Institute, Reino Unido), un programa que permite visualizar las lecturas mapeadas contra un genoma de referencia. A cada mutación se le asociará un valor de frecuencia, determinado por la siguiente ecuación:

$$\% \text{ Alelo mutado} = (\text{Lecturas mutación} / \text{Lecturas de la posición}) * 100 \quad (\text{Ec.1})$$

## 4. Resultados

### 4.1. Resultados obtenidos mediante PCR en tiempo real (plataforma Cobas®)

Del total de 40 muestras seleccionadas para el estudio, 14 (35,0%) de ellas, correspondientes a 10 pacientes, mostraron algún tipo de mutación en el gen *EGFR*. Las restantes 25 (62,5%)

mostraron un fenotipo *wild type* para el gen *EGFR*. De las 14 muestras mutadas, 10 (71,43%) mostraron una única mutación, mientras que las otras 4 (28,57%) corresponden a diversas mutaciones concurrentes con la p.T790M. En la Figura 5 se muestra la distribución de las mutaciones encontradas.

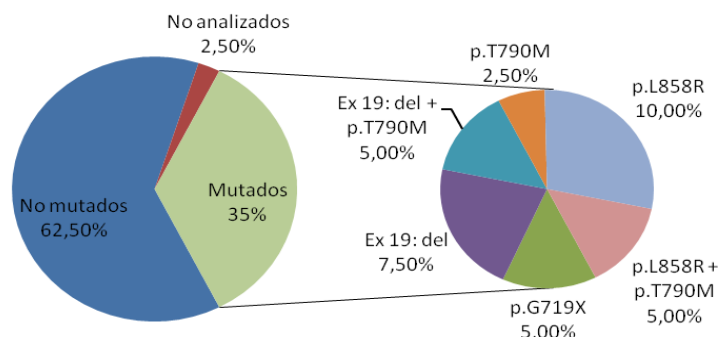


Figura 5: Gráfico de sectores de los resultados obtenidos mediante *Cobas Analyzer*. El primer gráfico muestra la distribución de muestras en función de si presentan mutación o no, o de si fueron analizadas. En el segundo gráfico se desglosan las mutaciones encontradas en las muestras positivas para formas alteradas de *EGFR*.

Las 2 mutaciones p.G719X identificadas en el exón 18 corresponden a una misma paciente de la que se disponen 2 biopsias diferentes. De las 5 mutaciones p.T790M del exón 20 (27,7% de las mutaciones identificadas), 4 de ellas iban acompañando a otra mutación: 2 casos de deleciones del exón 19, y los otros dos junto al cambio p.L858R del exón 21. De los 4 casos, 3 de ellos son segundas biopsias. Esto implica que en una primera biopsia se ha identificado la mutación no correspondiente al exón 20, detectándose en la segunda muestra la misma mutación junto a la mutación de resistencia p.T790M. Solo un caso muestra la p.T790M junto a otra mutación, en concreto una deleción del exón 19, habiendo exclusivamente una biopsia de este paciente.

Solo existe un caso que no haya sido analizado por cobas y que esté incluido en el estudio.

## 4.2.Confirmación de resultados mediante secuenciación directa

Las 14 muestras que mostraron alguna mutación al ser analizadas mediante Cobas fueron confirmadas mediante secuenciación directa. El total de reacciones de amplificación fue de 18: 2 del exón 18, 5 del exón 19, 5 del exón 20 y 6 del exón 21. Las muestras se organizaron en 7 tandas para ser lanzadas en el secuenciador capilar, habiendo en cada tanda un blanco y una muestra *wt* para la mutación ubicada en el exón analizado. Así, se secuenciaron 1 tanda de muestras del exón 18, 2 del exón 19, 2 del exón 20 y 2 del exón 21. Previamente a la electroforesis capilar, se testó mediante electroforesis en gel de agarosa la existencia de producto amplificado y la ausencia de contaminación (Figura 6). En todos los casos, la secuencia obtenida se daba por válida si la secuencia obtenida a partir del cebador directo y reverso concordaba.



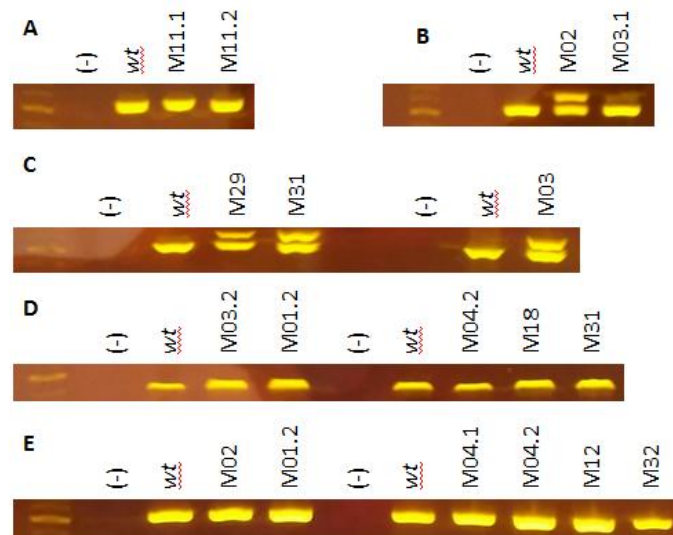


Figura 6: Captura con todas las bandas de los geles realizados para testar las PCRs destinadas a secuenciación directa. En todas las imágenes hay uno o dos negativos (-) que se ven claramente sin contaminación y el mismo número de muestras wt que muestran la amplificación de una secuencia sin mutar. (A) Bandas correspondientes al exón 18 de las muestras y el control. (B) y (C) Bandas correspondientes al exón 19 de las muestras y los controles. Se ve claramente en las muestras una doble banda, correspondiente a la delección y a la secuencia wild type. (D) Bandas correspondientes al exón 20 de los controles y las muestras. (E) Bandas correspondientes al exón 21 de los controles y las muestras.

Del total de los 18 exones amplificados, 17 mostraron (94,4%) alguna mutación. Todas las mutaciones detectadas para los exones amplificados vienen recogidas en la Tabla 3 y en la Figura 7. En las bandas producidas tras las electroforesis (Figura 6B y 6C) se ve una doble banda clara en aquellas muestras que han mostrado tener una delección. En el caso de la muestra M03.1 (Figura 6B), donde la mutación no ha sido identificada mediante esta técnica, la banda superior es mucho menos intensa en comparación que las de las otras muestras con mutación en el exón 19.

En relación con los resultados obtenidos con el Cobas, solo dos muestras no mostraron concordancia entre ambas técnicas. La primera de ellas (M18) corresponde a una muestra que mostró un cambio polimórfico en homocigosis en el exón 20, la cual resultó tener el cambio p.T790M según el análisis mediante Cobas y no mostrarlo en la secuenciación directa. Esta es la única muestra que posee el cambio polimórfico en homocigosis, habiéndose encontrado en heterocigosis en otras muestras analizadas que presentaban también la mutación p.T790M. El otro caso de no concordancia (M3.01) corresponde con una delección del exón 19, el cual fue identificado mediante Cobas, pero no mediante la secuenciación directa. En el caso de M03.1, la repetición del proceso desde la amplificación de la muestra dio el mismo resultado.

Al igual que con la tecnología Cobas, una de las muestras incluidas para su análisis no fue secuenciada mediante electroforesis capilar.





### 4.3. Estudios de secuenciación masiva mediante PGM

En total se ha realizado la secuenciación de 40 muestras empleando para ello 8 chips (Anexo II), en los que eran secuenciados simultáneamente el ADN y los transcritos de fusión de hasta 8 muestras. En total, se mapearon 15907996 lecturas de ADN con una media de 397700 lecturas por muestra, y 4841434 lecturas de ARN, con una media de 121036 lecturas por muestra. Según los criterios establecidos, dos muestras de ADN quedaron descartadas por no alcanzar el mínimo de lecturas requerido (100000), al igual que dos muestras de ARN (20000). Por tanto, se dispone de ADN válido de 38 muestras y ARN de 38 muestras, mientras que de ambas solo se dispone de resultados de 36 muestras.

#### 4.3.1. Mutaciones en el panel de ADN

En total, mediante PGM se han identificado un total de 102 mutaciones ubicadas en 13 genes diferentes: 19 en *EGFR* (18,62%), 1 en *HER2* (0,98%), 2 En *DDR2* (1,96%), 4 en *MET* (3,92%), 20 en *KRAS* (19,60%), 4 en *PIK3CA* (3,92%), 3 en *BRAF* (2,94%), 1 en *AKT1* (0,98%), 1 en *PTEN* (0,98%), 1 en *STK11* (0,98%), 1 en *CTNNB1* (0,98%) y 45 en *TP53* (44,11%). En el resto de los 10 (45,45%) genes del panel para ADN (*FBXW7*, *ERBB4*, *NOTCH 1*, *SMAD4*, *NRAS*, *MAP2K1*, *FGFR1*, *FGFR2*, *FGFR3* y *ALK*) no se ha identificado mutación alguna (Anexo III). En la Figura 9 se representa como se distribuyen las mutaciones en las diferentes muestras.

En el gen *EGFR* se han detectado mutaciones en 13 (32,5%) de las 40 muestras, pero dos de ellas no llegaban al número de lecturas establecido, de forma que han sido obviadas. Por tanto, solo se han detectado mutaciones en 12 (31,57%) pacientes de los 38 válidos. Tras el análisis, se han identificado un total de 19 mutaciones, compuestas por un total de 9 mutaciones diferentes: una vez las mutaciones p.E709K y p.G719C del exón 18; dos delecciones p.E746-A750del, 2 mutaciones puntuales (p.A750P y p.T751A) y una mutación compleja p.L747-P753delinsS del exón 19; los cambios p.T790M y p.C797S del exón 20; y la mutación p.L858R del exón 21 (Figura 8).

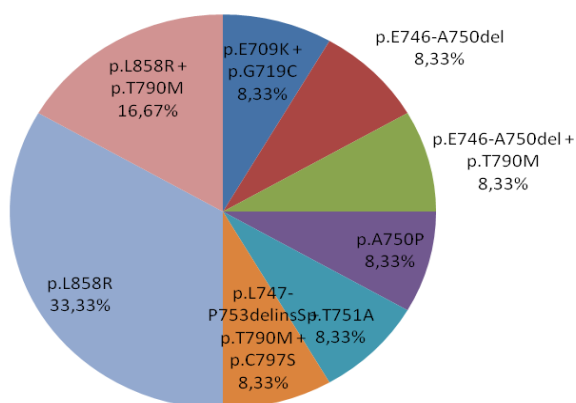
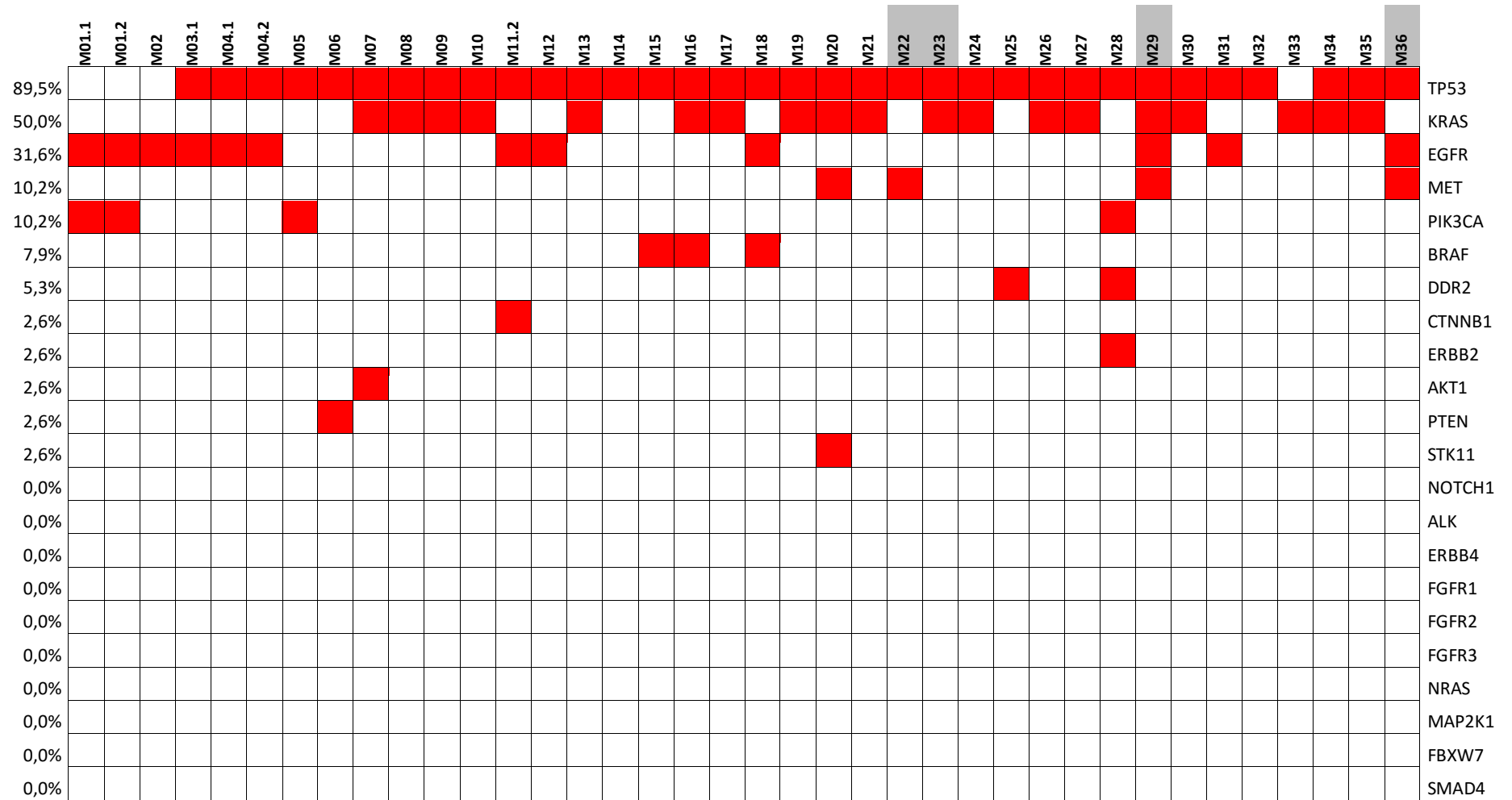


Figura 8: Distribución de las mutaciones identificadas en EGFR a lo largo de todas las muestras de NGS. Solo se incluyen muestras que hayan sido valorables (más de 100000 lecturas).

Figura 9: *heatmap* de las mutaciones de las muestras analizadas mediante PGM con el kit *Oncomine Solid Tumor ADN kit*. En gris las muestras de metástasis.



En el oncogén *KRAS* se han identificado un total de 20 mutaciones en 19 de las muestras (50,0%). En total, todas esas mutaciones afectan a 4 codones: en el codón 12 se han observado 8 cambios p.G12C y 6 correspondientes a p.G12V, en la posición 13 se han encontrado 2 cambios p.G13D y una mutación p.G13E, en el 61 se han registrado 2 cambios p.Q61H y en el 58 un único cambio p.T58I (Figura 10). Este último cambio se ha detectado simultáneamente con un cambio p.G12C. De las 19 muestras portadoras de al menos una mutación en *KRAS*, solo 1 (2,6%) mostraba simultáneamente mutaciones en el gen *KRAS* y *EGFR*. Esta muestra (M29) poseía las mutaciones p.G12C en *KRAS* y, según la NGS, la mutación p.A750P en *EGFR*. Además, la biopsia procede de una metástasis detectada en el hígado.

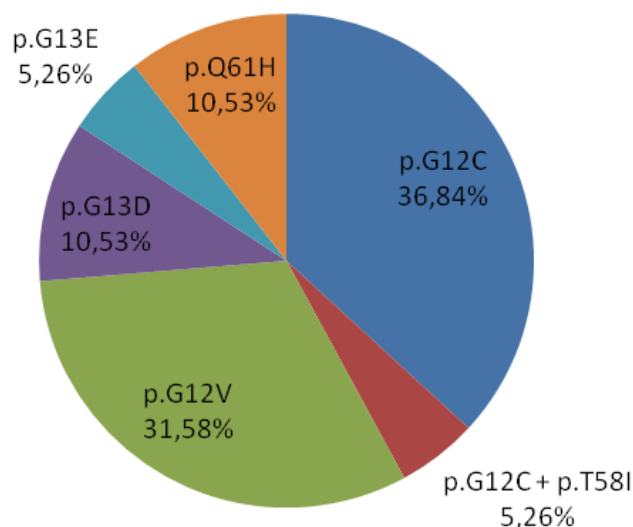


Figura 10: Distribución de las mutaciones localizadas en *KRAS* mediante NGS. Sólo se incluyen aquellas mutaciones de muestras validas.

En el gen *MET* se han detectado 3 mutaciones diferentes en 4 pacientes (10,52%). Los cambios identificados han sido p.E168D en dos muestras, p.M362T en una y p.N375S en otra. El cambio p.M362T se da un paciente que también posee la mutación p.G12V en *KRAS* y uno de los dos casos con p.G168E es la misma muestra que posee mutación en *KRAS* y *EGFR*, comentada en el párrafo anterior.

Mediante los filtros aplicados para el análisis de las mutaciones, en el gen *TP53* de las 38 muestras validas se ha detectado el polimorfismo p.P72R (c.215C>G) en 33 de ellas. Descontando esta variante, se han encontrado 12 muestras (31,57%) con mutación, todas ellas diferentes. De estas 12, solo una de ellas es un cambio no puntual siendo responsable de un *frame shift*.

En referencia al resto de genes, en *AKT1* se ha encontrado la mutación p.D46E, en *BRAF*, p.V600E y p.G469V; en *CTNNB1*, la p.S37C; en *DDR2*, p.R260L y p.M441I; en *ERBB2*, p.I767M; en *PIK3CA*, la p.E542Q, la p.G1049R y la p.H1047L; en *PTEN* un patrón de splicing alternativo; y en *STK11* un *frame shift* desde la posición p.P280.

### 4.3.2. Traslocaciones en el panel de ARN

Del total de 40 muestras para las que se analizaron los transcritos de fusión, solo 2 no han alcanzado las 20.000 lecturas mínimas, resultando no valorables. Doce de las 38 muestras válidas han mostrado poseer alguna alteración, verificada mediante el programa de análisis o no. Los dos casos verificados poseen reordenamientos del gen *ALK* con distintos puntos de rotura, M14 posee el transcrito *EML4(13)-ALK(20)* y M25 el *EML4(6)-ALK(20)*. Las restantes 10 muestras, son reordenamientos dudosos que están pendientes de confirmar por un método alternativo: 5 (13,15% de las muestras válidas) de ellas poseen reordenamientos en el gen *ROS1* y las restantes 5 (13,15% de las muestras válidas) en el gen *RET*, sin que el software indique el *partner* con quien se reordenan. Existen claros indicios de que estos 10 casos de reordenamientos dudosos sean negativos al coexistir con alteraciones somáticas en otros oncogenes del panel de ADN. Las muestras con reordenamientos confirmados en *ALK* no han mostrado mutaciones en genes como *EGFR* o *KRAS*. En cambio, ambas muestras poseen mutación en *TP53* y una de ellas posee un cambio en el gen *DDR2*.

### 4.4. Comparación de resultados entre los métodos moleculares empleados

De las 14 muestras que han sido analizadas mediante Cobas, secuenciación directa y secuenciación masiva, dos de ellas han sido obviadas debido a las bajas lecturas obtenidas en la secuenciación masiva, lo que las convierte en no valorables. De las 12 restantes, 7 muestras (58,3%) han obtenido en los tres análisis resultados concordantes (Tabla 3).

De los diferentes casos discordantes, el de la muestra M18 corresponde a una detección de la mutación p.T790M en el exón 20 mediante cobas, pero su no identificación mediante las otras dos tecnologías. En su lugar se identifica el polimorfismo c.2391G>A en homocigosis (p.Q787Q) mediante ambas tecnologías.

En otros dos casos (M02 y M29), tanto Cobas como la secuenciación directa identifican una deleción en el exón 19, mientras que por secuenciación masiva solo se detecta una mutación puntual en el codón 750 (M29) o 751 (M02). En ambos casos se examinó el archivo BAM generado tras la secuenciación mediante el programa *Integrative Genome Viewer* (IGV). En ambos casos se verificó la deleción correspondiente a cada secuencia y la ausencia de la mutación puntual que erróneamente había interpretado *Ion Reporter*. (Figura 11A y 11B).

Otro de los casos discordantes (M32) fue debido a que la NGS no fue capaz de detectar la mutación p.L858R, detectada mediante Cobas y secuenciación directa.

El último de los casos (M03.1) fue debido a que la secuenciación directa no detectó una mutación que tanto el Cobas como la NGS reconocieron. Posteriores repeticiones de la secuenciación directa de la muestra, tanto de la primera como de la segunda alícuota dieron el mismo resultado.

Tabla 3: Comparación de los resultados ofrecidos por las técnicas de análisis de EGFR. Cada línea corresponde a las mutaciones en un mismo exón. (MND – Mutación no detectada / (NV) - Muestra No Valorable / \* - Muestra con discordancias)

| Muestra | Exón | Resultado Cobas | Resultado Secuenciación directa              | Secuenciación masiva                         |              |
|---------|------|-----------------|--|--|--------------|
|         |      |                 |  | Resultado                                    | Mapped Reads |
| M01.1   | 21   | p.L858R         | p.L858R (c.2573 T>G)                         | p.L858R (c.2573 T>G)                         | 519,807      |
| M01.2   | 21   | p.L858R         | p.L858R (c.2573 T>G)                         | p.L858R (c.2573 T>G)                         | 554,289      |
|         | 20   | p.T790M         | p.T790M (c.2369 C>T)                         | p.T790M (c.2369 C>T)                         |              |
| M02*    | 19   | Ex19 del        | p.E746-T751delinsVA (c.2237-2253delinsTTGCT) | p.T751A (c.2251A>G)                          | 551,861      |
| M03.1*  | 19   | Ex19 del        | MND  | p.E746-A750del (c.2235-2249del)              | 518,364      |
| M03.2   | 19   | Ex19 del        | p.E746-A750del (c.2235-2249del)              | -  | 6,351 (NV)   |
|         | 20   | p.T790M         | p.T790M (c.2369 C>T)                         | -  |              |
| M04.1   | 21   | p.L858R         | p.L858R (c.2573 T>G)                         | p.L858R (c.2573 T>G)                         | 402,052      |
| M04.2   | 21   | p.L858R         | p.L858R (c.2573 T>G)                         | p.L858R (c.2573 T>G)                         | 503,515      |
|         | 20   | p.T790M         | p.T790M (c.2369 C>T)                         | p.T790M (c.2369 C>T)                         |              |
| M11.1   | 18   | p.G719X         | p.E709K (c.2125 G>A)<br>p.G719C (c.2155 G>T) | -  | 47,478 (NV)  |
| M11.2   | 18   | p.G719X         | p.E709K (c.2125 G>A)<br>p.G719C (c.2155 G>T) | p.E709K (c.2125 G>A)<br>p.G719C (c.2155 G>T) | 330,619      |
| M12     | 21   | p.L858R         | p.L858R (c.2573 T>G)                         | p.L858R (c.2573 T>G)                         | 303,224      |
| M18*    | 20   | p.T790M         | MND [p.Q787Q (c.2361G>A)]                    | MND [p.Q787Q (c.2361G>A)]                    | 431,259      |
| M29*    | 19   | Ex19 del        | p.L747-A750delinsP (c.2238-2248delinsGC)     | p.A750P (c.2248G>C)                          | 448,994      |
| M31     | 10   | Ex19 del        | p.E746-A750del (c.2235-2249del)              | p.E746-A750del (c.2235-2249del)              | 348,584      |
|         | 20   | p.T790M         | p.T790M (c.2369 C>T)                         | p.T790M (c.2369 C>T)                         |              |
| M32*    | 21   | p.L858R         | p.L858 (c.2573 T>G)                          | p.L858R (c.2573 T>G)                         | 300,795      |

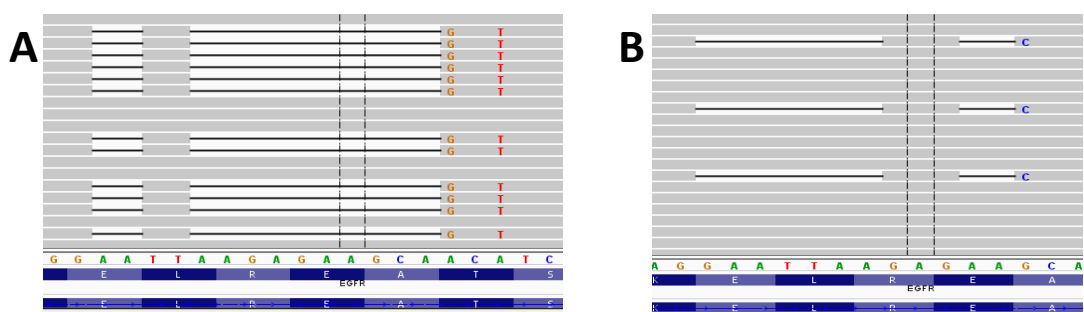


Figura 11: alineamientos de las lecturas obtenidas de dos de las muestras conflictivas. (A) – Alineamientos de la secuencia del exón 19 de *EGFR* en M02. Se ve claramente la presencia de una delección y del cambio ACA>GCT, este segundo reportado por el Ion Reporter. (B) – Alineamientos de la secuencia del exón 19 de *EGFR* en M51. Es posible ver la delección y el cambio reportado (GCA>CCA) en una proporción reducida frente a la secuencia wt.

En el caso válido M04.2 con mutación en el exón 18, la secuenciación masiva y la secuenciación directa identificaron una segunda mutación: p.E709K (c.2125G>A) que no era reportada por Cobas. Del mismo modo, en M04.1 (mapped reads<100000), del mismo paciente, la secuenciación capilar detecta esta misma segunda mutación.

## 5. Discusión

### 5.1.Comparación qPCR, secuenciación directa y NGS

El siguiente trabajo ha mostrado el desarrollo de tres aproximaciones para el estudio de las alteraciones del gen *EGFR*, importante para el tratamiento del CPNM (García-Foncillas *et al.*, 2011; Gómez *et al.*, 2012; Lindeman *et al.*, 2013). En concreto, este trabajo pretende comparar dos métodos habituales de trabajo implantados en el Laboratorio de Biología Molecular del HUyP La Fe, como son la secuenciación directa y la detección mediante qPCR, con una nueva metodología como es la *targeted sequencing* mediante NGS, que poco a poco va implantándose en la clínica. En los años recientes, con la aparición e implantación de este tipo de tecnologías, ya se han realizado estudios que pretenden mostrar como las NGS pueden igualar e incluso aportar más información que tecnologías de uso en la rutina clínica (Chin *et al.*, 2013; Tuononen *et al.*, 2013; Cernomaz *et al.*, 2016).

Hemos presentado un porcentaje de mutación del gen *EGFR* del 35,0% (14/40) muy superior a la prevalencia mutacional descrita para este gen en pacientes con CPNM. Es necesario destacar que este trabajo no pretende establecer las prevalencias mutacionales de cada una de las alteraciones descritas en pacientes con CPNM. Por el contrario intenta poner a punto un nuevo método de análisis, como es la secuenciación masiva, e incorporarlo en la rutina asistencial. Por ello, ha sido necesario seleccionar muestras de pacientes cuyos resultados eran previamente conocidos.

De las 12 muestras que presentaron mutaciones en *EGFR* mediante qPCR, se han encontrado discordancias con la secuenciación directa y con la NGS en 5 de ellas. Dos de ellas (M02 y M29) eran debidas a deleciones que a través de qPCR y de secuenciación directa fueron identificadas, pero la NGS no las reportaba. La revisión realizada en el IGV de ambas muestras (Figura 11 A y 11 B) muestra como la deleción sí está presente. La posible causa de que en ambas deleciones fueran reportadas como un cambio aminoacídico reside en el alineamiento realizado de la secuencia. Debido al funcionamiento del software *Ion Reporter*, cuando el programa analiza una posición concreta, intenta identificar una serie de variantes. Si no es capaz de encontrar alguna de las variantes que tiene predeterminadas, el programa no reporta que se haya producido una deleción. Debido a que en ambos casos aparecen varios nucleótidos alineados en mitad de la deleción, el programa no lo asocia a ninguna variante y no es capaz de reportar el cambio. Para futuros casos, será necesario proponer una mejora en el software de análisis que sea capaz de reconocer esta variación en el número de copias. En el caso concreto de *EGFR*, será importante una revisión manual de las secuencias generadas en

caso de que en el exón 19 se detecte un cambio que no sea una delección, la variación más prevalente en dicha región (Lindeman *et al.*, 2013).

Otro de los casos de discordancia ha sido la muestra M03.1, en la cual no se detectaba mediante secuenciación directa la mutación detectada por qPCR y NGS. Analizando los resultados de la NGS se ha identificado que la frecuencia del alelo mutado era del 3,03% (Anexo III). Debido a que tanto la qPCR como la NGS presentan una elevada sensibilidad, la mutación ha podido ser identificada. En cambio, la secuenciación directa no es capaz de detectar porcentajes tan bajos, estableciéndose su límite de detección en torno al 25% (García-Foncillas *et al.*, 2011; Gómez *et al.*, 2012). En un área como la oncología, y en concreto en muestras FFPE donde la población tumoral puede aparecer en muy baja proporción, la secuenciación directa puede ser fuente de falsos negativos, convirtiéndola en una tecnología susceptible de ser sustituida por otras de mayor sensibilidad. Por otro lado, también es importante la calidad de la muestra obtenida, en el caso de la M32, donde a partir de una misma extracción se habían generado dos eluciones, la mutación p.L858R solo se detectaba en la primera alícuota, evaluada tras obtener resultados negativos al analizar la segunda elución. En este sentido, siempre, la segunda elución va a ser de peor calidad, aunque en ocasiones es interesante disponer de este segundo volumen de muestra, la calidad puede comprometer los resultados de la prueba y llevar a falsos negativos.

La última discordancia corresponde a un individuo (M18) homocigoto para el polimorfismo c.2361G>C (p.Q787Q) que mediante Cobas se detectó la mutación p.T790M, mientras que la secuenciación directa y la NGS no reportaron el mismo resultado. A nivel de ensayo, el desconocimiento de la composición de la sonda impide establecer claramente el motivo del error en la qPCR, aunque es posible que sea responsable de una mala hibridación de la sonda, favoreciendo el falso positivo. Se requerirían otro tipo de ensayos para comprobar tal hipótesis. Este polimorfismo (rs1050171) es una mutación germinal de alta prevalencia en la población (Cernomaz *et al.*, 2016) que se encuentra tanto en personas sanas como en pacientes de CP, normalmente en heterocigosis. Un reciente estudio realizado por Koh *et al.* (2016) relaciona la presencia de esta mutación con un peor pronóstico en pacientes en etapas tempranas de CP de células escamosas.

Los resultados obtenidos en el resto de muestras han sido concordantes entre los tres ensayos. Cabe resaltar como en las muestras en las que existe una segunda biopsia (salvo M03.1 y M03.2) la mutación detectada en la primera muestra es también detectada en la segunda. Incluso en los casos donde los resultados de la NGS no son valorables, la mutación es detectada. Fuera de la coincidencia, realmente no se puede considerar como válido el resultado por incumplimiento del criterio de las lecturas. Es importante destacar la aparición de la mutación p.T790M en algunas de las segundas biopsias. Esta mutación está relacionada con las resistencias a los ITK de primera generación y es capaz de ser superada por los de tercera generación. Denominada mutación *gatekeeper*, aparece en pacientes que están siendo tratados con ITK de primer generación, siendo posible que prevalezca por selección positiva (Gazdar, 2009). Apareciendo inicialmente en baja proporción, los test de baja sensibilidad no serían capaces de detectarla hasta que el tratamiento hace aumentar la población tumoral con la resistencia. Una vez más se destaca la necesidad de métodos



sensibles que nos permitan detectar las mutaciones en sus concentraciones más bajas. Actualmente, se está trabajando en métodos que permitan aumentar la sensibilidad (Zhao *et al.*, 2016), así como que permitan detectar la mutación en otro tipo de muestras, como la biopsia líquida. Estas metodologías requieren mayor sensibilidad debido a la baja cantidad de circulating tumor ADN (ctDNA) que hay en la sangre (Murtaza *et al.*, 2013; Jiang *et al.*, 2015; Thress *et al.*, 2015).

## 5.2. Impacto de la tecnología de NGS

Mientras que la qPCR es un ensayo limitado por la especificidad de las sondas empleadas y la secuenciación directa por los amplicones seleccionados, la NGS está menos limitada en este sentido por la capacidad de multiplexar diversos amplicones que nos proporcionaran una información integral y más profunda del tumor a analizar. Por su parte, con una sonda de qPCR podemos testar una mutación o conjunto de mutaciones muy concretas, a veces sin ni siquiera conocer qué mutación estamos detectando. En cambio, las tecnologías de secuenciación permiten detallar cualquier variación en una secuencia, siempre y cuando esté en la suficiente cantidad para poder ser identificada por los detectores del aparato. En el caso de la secuenciación directa, esta sensibilidad es baja y en una misma electroforesis el número de amplicones analizados simultáneamente dependerá del número de capilares disponibles en el aparato. En la secuenciación masiva este inconveniente desaparece; la elevada sensibilidad de la tecnología y el poco requerimiento de muestra permiten que con poca cantidad de ADN se pueda detectar con facilidad cualquier mutación. Además, la naturaleza del proceso de generación de las librerías permite, por un lado, multiplexar el ensayo a través del uso de adaptadores identificativos, y por otro, analizar varios amplicones mediante el uso de varios pares de cebadores, que gracias a su posterior alineamiento, se pueden analizar a la vez sin miedo de que sean confundidos con regiones de secuencias que no es la correspondiente. Por estos motivos, la NGS se presenta como una tecnología potente desde la que un único ensayo nos permite obtener mucha más información que las otras tecnologías a la vez que alcanza una elevada sensibilidad. Así, en este estudio, además de las mutaciones existentes en *EGFR*, se han podido detectar mutaciones en otros 11 genes, así como traslocaciones y reordenamientos.

Las mutaciones en el gen *EGFR* se han encontrado en un 31,6% de las muestras seleccionadas. Las deleciones en el exón 19 son consideradas mutaciones de clase I, mientras que las mutaciones puntuales como p.G719C o p.L858R son de clase II. Ambos tipos de mutaciones son consideradas activantes y afectan al dominio TK, generando un EGFR con mayor afinidad por el ATP. Estas mutaciones actúan como marcadores predictivos de respuesta a ITK, de forma que aquellos pacientes con estas mutaciones se podrán beneficiar de ellos (Gazdar, 2009). Es importante resaltar como en segundas biopsias de pacientes con estas mutaciones activantes (M01.2 y M11.2), han acabado manifestando la mutación p.T790M surgida como resistencia a estos tratamientos (Garraway y Jänne, 2012; González de Castor *et al.*, 2015). Pero al igual que hay fármacos útiles que permiten superar esta mutación (ITKs de tercera generación), ya han surgido casos en los que han aparecido resistencias. Es el caso de la muestra de metástasis hepática M36, en el que se observa además la presencia de la mutación

p.C797S. La pérdida del grupo sulfuro de la cisteína impide la formación del enlace covalente de los ITK de 3ª generación, perdiendo completamente su eficacia (Yu et al., 2015). Además de este mecanismo, como se ha comentado en la introducción, se han reportado casos que emplean otros mecanismos de resistencia como la amplificación de *MET* (Planchard et al., 2015).

En *KRAS* se han detectado 20 mutaciones distribuidas entre 19 pacientes, 19 de las mutaciones, situadas en las posiciones 12, 13 y 61, favorecen la pérdida de la actividad GTPasa de *KRAS*. La otra mutación, p.T58I presenta una baja frecuencia, habiendo sido encontrado en algunos estudios (Goranova et al., 2011). Esta mutación, cuando es germinal, se asocia a otras patologías como el síndrome de Noonan (Schubbert et al., 2006). En comparación con *EGFR*, del total de 40 muestras analizadas, 32 (80,0%) mostraban al menos una mutación en *KRAS* o en *EGFR*. De todas ellas, solo la muestra M29 mostraba mutaciones en ambos, siendo la p.L747-A750delinsP en *EGFR* y la p.G12C en *KRAS*. Estos casos son de muy baja frecuencia y actualmente existe información limitada sobre el efecto de ambas mutaciones en el tratamiento con ITK contra *EGFR*, aunque se cree que la presencia de ambas mutaciones se debe a una presencia multiclonal, haciendo susceptibles a unos clones de los tratamientos de ITK dirigidos contra *EGFR* y a otros no (Benesova et al., 2010). En el estudio de Ulivi et al. (2015), tres pacientes con CP de células escamosas fueron *EGFR* y *KRAS* positivos, dos de los cuales fueron tratados con gefitinib, mostrando una situación estable de la enfermedad.

*TP53* es un TSG encargado del mantenimiento del ciclo celular. Un fallo en cualquiera de sus subunidades conlleva al error de toda la proteína. En total, se han identificado 50 mutaciones en las 38 muestras, agrupadas en 13 tipos de mutaciones. En 33 de las muestras se ha detectado el polimorfismo p.P72R. Esta variante ha sido muy estudiada, y aunque asocian la presencia de p.R72 con una mejor apoptosis que con el alelo p.P72, no se ha identificado una relación significativa ni con el CP ni con otros cánceres (Olivier et al., 2010). Respecto a las otras 12 mutaciones, la mitad de ellas se han encontrado en *hotspots* del dominio de unión al ADN definidos en CP: p.V157F, p.R158L, p.R248L, p.249S y p.R273L (Gibbons et al., 2014).

En el gen *BRAF*, las mutaciones encontradas son de carácter patogénico. En concreto, la mutación p.V600E favorece la constante actividad de *BRAF* mediante la introducción del grupo carboxilo. Este molécula simula las modificaciones electroquímicas inducidas por un grupo fosfato, imitando las señales activadoras y favoreciendo la actividad constitutiva (Sanchez-Torres et al., 2013; Candarella et al., 2015). La otra mutación, p.G469V también es considerada una mutación activante, aunque se presenta con menos frecuencia (Candarella et al., 2015).

*PIK3CA* codifica para la subunidad catalítica de la proteína PI3K. Las mutaciones encontradas en los 4 de los pacientes (Figura 9) han sido estudiadas en cáncer de colon y son consideradas mutaciones activantes. Las mutaciones de los codones 542 y 545 afectan al exón 9, que codifica para parte de un dominio helicoidal, mientras que las que alteran los codones 1047 y 1049 afectan directamente al dominio tirosina cinasa. De las 4, las mutaciones en la posición 1049 son las menos frecuentes (Liao et al., 2013).

En *STK11*, una serin-tirosin cinasa, se ha detectado la delección de una citosina en la posición 837 (c.837-837del) que es responsable de un *frame shift* en la proteína, generando un codón

de stop en la posición p.286. Esta mutación es considerada como inactivante. Cuando fue descrita por Sanchez-Céspedes *et al.* (2001), la mutación aparecía junto a otra en *KRAS*, como ocurre en la muestra M20.

El resto de mutaciones detectadas por la NGS se corresponden con polimorfismos o no han sido descritas en CP previamente. Así, en el gen *MET* las tres variantes detectadas se corresponden con polimorfismos (Petrini *et al.*, 2015). En cambio, entre variables no identificadas en CP encontramos p.D46E en *AKT1*; p.S37C en *CTNNB1*, reportada en cáncer de colon (Mikami, 2006); p.R260L y p.M441I en *DDR2*; p.I767M en *ERBB2*, estando esta última descrita en cáncer de mama (Chmielecki *et al.*, 2015); o el splicing alternativo generado en *PTEN*.

En el caso de los transcritos de fusión, de las 38 muestras válidas solo se ha podido identificar reordenamiento claramente en dos de las muestras. En ambos casos ha sido para el reordenamiento *EML4-ALK*, en concreto *EML4(6)-ALK(20)* en M25 y *EML4(13)-ALK(20)* en M14. Los reordenamientos en este gen presentan una baja frecuencia en los pacientes de CPNM, pero aumenta en pacientes no fumadores y de adenocarcinomas (Gómez *et al.*, 2012). En 10 casos, la secuenciación masiva ha mostrado resultados dudosos para reordenamientos en *ROS1* y *RET* siendo además incapaz de mostrar el *partner* con el que reordenaron. En estos casos se requeriría una mejora del proceso de identificación de estos transcritos, para evitar tener que recurrir a pruebas complementarias para la confirmación de dichos reordenamientos.

Pese a toda la información que aporta la secuenciación masiva, no toda ella es de utilidad clínica. Además, requiere personal cualificado y preparado para los largos protocolos que plantea y en los que se trabaja con volúmenes pequeños y con poca muestra. A eso hay que sumar el elevado coste de la reacción y de los equipos. A nivel de rutina clínica, en ciertas situaciones puede ser interesante emplear pruebas que requieran de menos tiempo y que sean más sencillas de llevar a cabo. Sin embargo, no cabe duda que el empleo de paneles diseñados frente a genes para los que se dispone de tratamientos específicos, puede ser de gran utilidad en el manejo de los pacientes, ayudando al clínico a seleccionar la opción terapéutica más adecuada a cada caso.

## 6. Conclusiones

En los anteriores apartados se ha procedido a comparar tres de los métodos actualmente empleados para la identificación de biomarcadores diagnósticos en CPNM: la qPCR a través de la plataforma cobas, la secuenciación directa y la secuenciación masiva mediante semiconductores. Con todo lo expuesto en las páginas anteriores, podemos concluir que:

- Aunque la secuenciación directa sea un método accesible para todos los laboratorios, tecnologías como la qPCR o la secuenciación masiva van desplazándola poco a poco. Debido a su baja sensibilidad, mediante secuenciación directa es posible encontrarnos con falsos negativos que priven al paciente de un tratamiento adecuado.

- Igualmente, la qPCR está limitada a sondas que detecten mutaciones específicas. Esto limita la capacidad de trabajo de estas técnicas, eficaces ante muestras positivas pero incompletas en el caso de las muestras negativas. La NGS supera este problema al ser capaz de analizar un mayor número de amplicones y genes en cada reacción. Además, como se ha comprobado, mutaciones próximas a las sondas pueden llevar a problemas en la hibridación de estas, llevando a la técnica a resultados erróneos.
- La secuenciación masiva ofrece la posibilidad de reducir el número de ensayos necesarios para obtener la misma información. Así, mediante la selección adecuada de los genes, podremos determinar todas las mutaciones de interés clínico con una única reacción de secuenciación y mínima cantidad de muestra, en vez de tener que emplear una prueba distinta para cada gen, y a veces, para cada fragmento del gen.
- Además de la identificación de biomarcadores, la secuenciación masiva puede aportarnos información de nuevas alteraciones. Por ello, existe la necesidad de aprender a interpretar y caracterizar toda la información que nos aporta. Por otra parte, la mala interpretación de los datos o un incorrecto análisis puede llevar a errores en los resultados obtenidos. Esto vuelve a indicar la necesidad de manejar con eficacia estas plataformas desde centros de reconocida trayectoria en biología molecular.

## 7. Referencias

### 7.1.Bibliografía

- BENESOVA, L.; MINARIK, M.; JANCARIKOVA, D.; BELSANNOVA, B.; and PESEK, M. (2010) Multiplicity of EGFR and KRAS Mutations in Non-small Cell Lung Cancer (NSCLC) Patients Treated with Tyrosine Kinase Inhibitors. *Anticancer Research*, 30: 1667-1672
- BIANKIN, A.; PIANTADOSI, S.; HOLLINGSWORTH, S. (2015) Patient-centric trials for therapeutic development in precision oncology. *Nature*, 526: 361-370
- BIOMARKER DEFINITIONS WORKING GROUP (2001) Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework. *Clin. Pharmacol. Ther.*, 69: 89-95
- BRAMBILLA, E.; GAZDAR, A. (2009) Pathogenesis of lung cancer signaling pathways: roadmap for therapies. *Eur Respir J*, 33: 1485–1497
- BUBNOFF, A. (2008) Next Generation Sequencing: The race is on. *Cell*, 132(5): 721-723
- CARDARELLA, S.; OGINO, A.; NISHINO, M.; BUTANEY, M.; SHEN, J.; LYDON, C.; YEAP, B. Y.; SHOLL, L. M.; JOHNSON, B. E.; JEANNE, P. A. (2015) Clinical, Pathologic, and Biologic Features Associated with BRAF Mutations in Non–Small Cell Lung Cancer. *Clin Cancer Res*; 19(16): 4532-4540
- CERNOMAZ, A. T.; MACOVEI, I. I.; PAVEL, I.; GRIGORIU, C.; MARINCA, M.; BATY, F.; PETER, S.; ZONDA, R.; BRUTSCHE, M.; GRIGORIU, B. D. (2016) Comparison of next generation sequencing, SNaPshot assay and real-time polymerase chain reaction for lung adenocarcinoma *EGFR* mutation assessment. *BMC Pulmonary Medicine*, 16(88): 1-7
- CHIN, E. L.; SILVA, C.; HEGDE, M. (2013) Assessment of clinical analytical sensitivity and specificity of next-generation sequencing for detection of simple and complex mutations. *BMC Genetics*, 14(6): 1-13
- CHMIELECKI, J.; ROSS, J. S.; WANG, K.; FRAMPTON, G. M.; PALMER, G. A.; ALI, S. M.; PALMA, N.; MOROSINI, D.; MILLER, V. A.; YELENSKY, R.; LIPSON, D.; STEPHENS, P. J. (2015) Oncogenic Alterations in ERBB2/HER2 Represent Potential Therapeutic Targets Across Tumors From Diverse Anatomic Sites of Origin. *The Oncologist*, 20: 7–12
- CROSS, D. A. E.; ASHTON, S. E.; GHIORGHIOU, S.; EBERLEIN, C.; NEBHAN, C. A.; SPITZLER, P. J.; ORME, J. P.; FINLAY, M. R. V.; WARD, R. A.; MELLOR, M. J.; HUGHES, G.; RAHI, A.; JACOBS, V. N.; BREWER, M. R.; ICHIHARA, E.; SUN, J.; JIN, H.; BALLARD, P.; AL-KADHIMI, K.; ROWLINSON, R.; KLINOWSKA, T.; RICHMOND, G. H. P.; CANTARINI, M.; KIM, D. W.; RANSON, M. R.; PAO, W. (2014) AZD9291, an Irreversible EGFR TKI, Overcomes T790M-Mediated Resistance to EGFR Inhibitors in Lung Cancer. *Cancer Discovery*, 4: 1046-1061.

- CUI, J. J.; TRAN-DUBÉ, M.; SHEN, H.; NAMBU, M.; KUNG, P.; PAIRISH, M.; JIA, L.; MENG, J.; FUNK, L.; BOTROUS, I.; MCTIGUE, M.; GRODSKY, N.; RYAN, K.; PADRIQUE, E.; ALTON, G.; TIMOFEEVSKI, S.; YAMAZAKI, S.; LI, Q.; ZOU, H.; CHRISTENSEN, J.; MROCZKOWSKI, B.; BENDER, S.; KANIA, R. S.; EDWARDS, M. P. (2011) Structure Based Drug Design of Crizotinib (PF-02341066), a Potent and Selective Dual Inhibitor of Mesenchymal Epithelial Transition Factor (c-MET) Kinase and Anaplastic Lymphoma Kinase (ALK). *J. Med. Chem.*, 54: 6342–6363
- DOEBELE, R. C.; PILLING, A. B; AISNER, D. L.; KUTATELADZE, T. G.; LE, A. T.; WEICKHARDT, A. J.; KONDO, K. L.; LINDERMAN, D. J.; HEASLEY, L. E.; FRANKLIN, W. A.; VARELLA-GARCIA, M.; CAMIDGE, D. R. (2012) Mechanisms of Resistance to Crizotinib in Patients with ALK Gene Rearranged Non–Small Cell Lung Cancer. *Clin Cancer Res*, 18(5): 1472-1482
- FERLAY, J.; SOERJOMATARAM, I.; DIKSHIT, R.; ESER, S.; MATHERS, C.; REBELO, M.; PARKIN, D. M.; FORMAN, D.; BRAY, F. (2015) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer*, 136: E359–E386
- GARCÍA-FONCILLAS, J.; GARRIDO, P.; GÓMEZ, J.; PALACIOS, J.; TARÓN, M. (2011) Recomendaciones para la determinación de las mutaciones del gen EGFR en el carcinoma de pulmón no microcítico. *Revista Española de Patología*, 44(1): 17-31
- GARRAWAY, L. A.; JÄNNE, P. A. (2012) Circumventing Cancer Drug Resistance in the Era of Personalized Medicine. *Cancer Discovery*, 2: 214-226
- GAZDAR, A. F. (2009) Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. *Oncogene*, 28: S24–S31
- GIBBONS, D. L.; BYERS, L. A.; KURIE, J. M. (2014) Smoking, p53 Mutation, and Lung Cancer. *Mol Cancer Res*; 12(1): 3-13
- GONZÁLEZ DE CASTRO, D.; CLARKE, P. A.; AL-LAZIKANI, B.; WORMAN, P. (2013) Personalized Cancer Medicine: Molecular Diagnosis, Predictive Biomarkers, and Drug Resistance. *Clinical Pharmacology & Therapeutics*, 93 (3): 252-259
- GORANOVA, T. E.; OHUE, M.; SHIMOHARU, Y.; KIKUYA KATO, K. (2011) Dynamics of cancer cell subpopulations in primary and metastatic colorectal tumors. *Clin Exp Metastasis*. (28): 427–435
- GREULICH, H. (2010) The Genomics of Lung Adenocarcinoma: Opportunities for Targeted Therapies. *Genes Cancer*, 1(12): 1200–1210.
- HANAHAN, D; WEINBERG, R. (2011) Hallmarks of Cancer: The Next Generation. *Cell*, 144: 646-674.
- HARRIS, R. C.; CHUNG, E.; COFFEY, R. J. (2003) EGF receptor ligands. *Experimental Cell Research*, 284: 2–13
- HENRY, N. L.; HAYES, D. F. (2012) Cancer biomarkers. *Molecular Oncology*, 6: 140-146

HERBST, R. S. (2004) Review of epidermal growth factor receptor biology. *Int. J. Radiation Oncology Biol. Phys.*, 59(2):21–26, 2004

HOVELSON, D. H.; MCDANIEL, A. S.; CANI, A. K.; JOHNSON, B.; RHODES, K.; WILLIAMS, P. D.; BANDLA, S.; BIEN, G.; CHOPPA, P.; HYLAND, F.; GOTTIMUKKALA, R.; LIU, G.; MANIVANNAN, M.; SCHAGEMAN, J.; BALLESTEROS-VILLAGRANA, E.; GRASSO, C. S.; QUIST, M. J.; YADATI, V.; AMIN, A.; SIDDIQUI, J.; BETZ, B. L.; KNUDSEN, K. E.; COONEY, K. A.; FENG, F. Y.; ROH, M. H.; NELSON, P. S.; LIU, C. J.; BEER, D. G.; WYNGAARD, P.; CHINNAIYAN, A. M.; SADIS, S.; RHODES, D. R.; TOMLINS, S. A. (2015) Development and Validation of a Scalable Next-Generation Sequencing System for Assessing Relevant Somatic Variants in Solid Tumors. *Neoplasia*, 17(4): 385-399

INGELMAN-SUNDBERG, M. (2015) Personalized medicine into the next generation. *Journal of Internal Medicine*, 277; 152–154

JIANG, T.; REN, S.; ZHOU, C. (2015) Role of circulating-tumor DNA analysis in non-small cell lung cancer. *Lung Cancer*, 90: 128–134

KOH, Y. W.; KIM, H. J.; KWON, H. Y.; HAN, J. H.; LEE, C. K.; LEE, M. S.; KIM, C. J.; BAEK, M. J.; JEONG, D. (2016) Q787Q EGFR Polymorphism as a Prognostic Factor for Lung Squamous Cell Carcinoma. *Clinical Translational Research* , DOI: 10.1159/000444495

LIAO, X.; MORIKAWA, T.; LOCHHEAD, P.; IMAMURA, Y.; KUCHIBA, A.; YAMAUCHI, M.; NOSHO, K.; QIAN, Z. R.; NISHIHARA, R.; MEYERHARDT, J. A.; FUCHS, C. S.; OGINO, S. (2012) Prognostic Role of PIK3CA Mutation in Colorectal Cancer: Cohort Study and Literature Review. *Clin Cancer Res*, 18(8): 2257–2268.

LINDEMAN, N. I.; CAGLE, P. T.; BEASLEY, M. B.; CHITALE, D. A.; DACIC, S.; GIACCONE, G.; JENKINS, R. B.; DAVID J. KWIATKOWSKI, D. J.; SALDIVAR, J. S.; SQUIRE, J.; THUNNISSEN, E.; LADANYI, M. (2013) Molecular Testing Guideline for Selection of Lung Cancer Patients for EGFR and ALK Tyrosine Kinase Inhibitors. *The Journal of Molecular Diagnosis*, 15(4): 415-453

LUO S. Y.; LAM, D. C. (2013) Oncogenic driver mutations in lung cancer. *Translational Respiratory Medicine*, 1(6): 1-8

MIKAMI, M.; NOSHO, K.; YAMAMOTO, H.; TAKAHASHI, T.; MAEHATA, T.; TANIGUCHI, H.; ADACHI, Y.; IMAMURA, A.; FUJITA, M.; HOSOKAWA, M.; ITOH, F.; IMAI, K.; SHINOMURA, Y. (2006) Mutational analysis of b-catenin and the RAS-RAF signalling pathway in early flat-type colorectal tumours. *European Journal Of Cancer*, 42: 3065–3072

MOK, T. S.; WU, Y. L.; THONGPRASERT, S.; YANG, C. H.; CHU, D. T.; SAIJO, N.; SUNPAWERAVONG, P.; HAN, B.; MARGONO, B.; ICHINOSE, Y.; NISHIWAKI, Y.; OHE, Y.; YANG, J. J.; CHEWASKULYONG, B.; JIANG, H.; DUFFIELD, E. L.; WATKINS, C. L.; ARMOUR, A. A.; FUKUOKA, M. (2009) Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma. *N. Eng J. Med.*, 361(10): 947-957

MURTAZA, M.; DAWSON, S. J.; TSUI, D. W. Y.; GALE, D.; FORSHEW, T.; PISKORZ, A. M.; PARKINSON, C.; CHIN, S. F.; KINGSBURY, Z.; WONG, A. S. C.; MARASS, F.; HUMPHRAY, S.; HADFIELD, J.; BENTLEY, D.; TAN MIN CHIN, T. M.; BRENTON, J. D.; CALDAS, C.; ROSENFELD, N.

(2013) Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature*, 497: 108–112

OLIVIER, M.; HOLLSTEIN, M.; HAINAUT, P. (2010) TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb Perspect Biol.*, 2(1): 1-13

PETRINI, I. (2015) Biology of MET: a double life between normal tissue repair and tumor progression. *Ann Transl Med*, 3(6): 82-96

PLANCHARD, D.; LORIOT, Y.; ANDRÉ, F.; GOBERT, A.; AUGER, N.; LACROIX, L.; SORIA J. C. (2015) EGFR-independent mechanisms of acquired resistance to AZD9291 in EGFR T790M-positive NSCLC patients. *Annals of Oncology*, 26: 2073–2078

RAPHAEL, B. J.; DOBSON, J. R.; OESPER, L.; VANDIN, F. (2014) Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine. *Genome Medicine*, 6(5): 1-17

RHODES, D. R.; KALYANA-SUNDARAM, S.; MAHAVISNO, V.; VARAMBALLY, R.; JIANJUN YU, J.; BRIGGS, B. B.; BARRETTE, T. R.; ANSTET, M. J.; KINCEAD-BEAL, C.; KULKARNI, P.; VARAMBALLY, S.; GHOSH, D.; CHINNAIYAN, A. M. (2007) Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Gene Expression Profiles. *Neoplasia*, 9 (2): 166-180

SÁNCHEZ-CESPEDES, M.; PARRELLA, P.; ESTELLER, M.; NOMOTO, S.; TRINK, B.; ENGLES, J. M.; WESTRA, W. H.; HERMAN, J. G.; SIDRANSKY, D. (2001) Inactivation of LKB1/STK11 Is a Common Event in Adenocarcinomas of the Lung. *Cancer Research*, 62: 3659–3662

SÁNCHEZ-TORRES, J. M.; VITERI, S.; MOLINA, M. A.; ROSELL, R. (2013) BRAF mutant non-small cell lung cancer and treatment with BRAF inhibitors. *Transl Lung Cancer Res*, 2(3): 244-250

SCALTRITI, M.; BASELGA, J. (2006) The Epidermal Growth Factor Receptor Pathway: A Model for Targeted Therapy. *Clin Cancer Res*, 12(18): 5268-5272

SCHUBBERT, S.; ZENKER, M.; ROWE, S. L.; BÖLL, S.; KLEIN, C.; BOLLAG, B.; VAN DER BURGT, I.; MUSANTE, L.; KALSCHUEER, V.; WEHNER, L. E.; NGUYEN, H.; WEST, B.; ZHANG, K. Y. J.; SISTERMANS, E.; RAUCH, A.; NIEMEYER, C. M.; SHANNON, K.; KRATZ, C. P. (2006) Germline KRAS mutations cause Noonan syndrome. *Nature Genetics*, 38(3): 331-336

SODA, M.; CHOI, Y. L.; ENOMOTO, M.; TAKADA, S.; YAMASHITA, Y.; ISHIKAWA, S.; FUJIWARA, S.; WATANABE, H.; KURASHINA, K.; HATANAKA, H.; BANDO, M.; OHNO, S.; ISHIKAWA, Y.; ABURATANI, H.; NIKI, T.; SOHARA, Y.; SUGIYAMA, Y.; MANO, H. (2007) Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. *Nature*, 448(7153): 561-566

SOLCA, F.; DAHL, G.; ZOEPHEL, A.; BADER, G.; SANDERSON, M.; KLEIN, C.; KRAEMER, O.; HIMMELSBACH, F.; HAAKSMA, E.; ADOLF, G. R. (2012) Target Binding Properties and Cellular Activity of Afatinib (BIBW 2992), an Irreversible ErbB Family Blocker. *THE JOURNAL OF PHARMACOLOGY AND EXPERIMENTAL THERAPEUTICS*, 343 (2): 342-350



STAMOS, J.; SLIWKOWSKI, M. X.; EIGENBROT, C. (2002) Structure of the Epidermal Growth Factor Receptor Kinase Domain Alone and in Complex with a 4-Anilinoquinazoline Inhibitor. *The Journal Of Biological Chemistry*, 277 (48): 46265–46272

THRESS, K. S.; BRANT, R.; CARR, T. H.; DEARDEN, S.; JENKINS, S.; BROWN, H.; HAMMETT, T.; CANTARINI, M.; BARRETT, J. C. (2015) EGFR mutation detection in ctDNA from NSCLC patient plasma: Across-platform comparison of leading technologies to support the clinical development of AZD9291. *Lung Cancer*, 90: 509–515

TUCKER, T.; MARRA, M.; FRIEDMAN, J. M. (2009) Massively Parallel Sequencing: The Next Big Thing in Genetic Medicine. *The American Journal of Human Genetics*, 85: 142–154

TUONONEN, K.; SARHADI, V. K.; WIRTANEN, A.; RÖNTY, M.; SALMENKIVI, K.; KNUUTILA, A.; REMES, S.; TELARANTA-KEERIE, A. I.; BLOOR, S.; ELLONEN, P.; KNUUTILA, S. (2013) Targeted Resequencing Reveals ALK Fusions in Non-Small Cell Lung Carcinomas Detected by FISH, Immunohistochemistry, and Real-Time RT-PCR: A Comparison of Four Methods. *Biomed Research International*, 2013 (757490): 1-9

ULIVI, P.; CHIADINI, E.; DAZZI, C.; DUBINI, A.; COSTANTINI, M.; MEDRI, L.; PUC CETTI, M.; CAPELLI, L.; CALISTRI, D.; VERLICCHI, A.; GAMBONI, A.; PAPI, M.; MARIOTTI, M.; DE LUIGI, N.; SCARPI, E.; BRAVACCINI, S.; TUROLLA, G. M.; AMADORI, D.; CRINÒ, L.; DELMONTE, A. (2015) Non-squamous, Non-Small-Cell Lung Cancer Patients Who Carry a Double Mutation of EGFR, EML4-ALK or KRAS: Frequency, Clinical Pathological Characteristics, and Response to Therapy. *Clinical Lung Cancer*, 2015: 1-7

VOGELSTEIN, B.; PAPADOPOULOS, N.; VELCULESCU, V. E.; ZHOU, S.; DIAZ, L.A.; KINZLER, K. W. (2013) Cancer Genome Landscape. *Science*, 339: 1546-1558

WAKELING, A. E.; GUY, S. P.; WOODBURN, J. R.; ASHTON, S. E.; CURRY, B. J.; BARKER, A. J.; GIBSON, K. H. (2002) ZD1839 (Iressa): An Orally Active Inhibitor of Epidermal Growth Factor Signaling with Potential for Cancer Therapy. *Cancer Research*, 62: 5749–5754

YARDEN, Y.; SLIWKOWSKI, M. X. (2001) Untangling the ErbB Signaling Network. *Nature reviews*, 2: 127-137

YU, H. A.; TIAN, S. K.; DRILON, A. E.; BORSU, L.; RIELY, G. J.; ARCILA, M. E.; LADANYI, M. (2015) Acquired Resistance of EGFR-Mutant Lung Cancer to a T790M-Specific EGFR Inhibitor: Emergence of a Third Mutation (C797S) in the EGFR Tyrosine Kinase Domain. *JAMA Oncology*, 1(7): 982-984

ZHAO, J.; FENG, H. H.; ZHAO, J. Y.; LIU, L. C.; XIE, F. F.; XU, Y.; CHEN, M. J.; ZHONG, W.; LI, L. Y.; WANG, H. P.; ZHANG, L. I.; XIAO, Y. I.; CHEN, W. J.; WANG, M. Z. (2016) A sensitive and practical method to detect the T790M mutation in the epidermal growth factor receptor. *Oncology Letters*, 11: 2573-2679

ZHOU, C.; WU, Y. L.; CHEN, G.; FENG, J.; LIU, X. Q.; WANG, C.; ZHANG, S.; WANG, J.; ZHOU, S.; REN, S.; LU, L.; ZHANG, L.; HU, C.; HU, C.; LUO, Y.; CHEN, L.; YE, M.; HUANG, J.; ZHI, X.; ZHANG, Y.; XIU, Q.; MA, J.; ZHANG, L.; YOU, C. (2011) Erlotinib versus chemotherapy as first-line

treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomized, phase 3 study. *Lancet Oncol*, 12: 735–742

## **7.2.Recursos Web**

THERMO FISHER SCIENTIFICS. Ion Torrent™ next-gen sequencing technology. Consultado el 6 de junio de 2016. Disponible en: <https://www.youtube.com/watch?v=WYBzbxlfuKs>

WORLD HEALTH ORGANIZATION (WHO) *Cancer*. Consultado el 5 de junio de 2016. Disponible en: <http://www.who.int/mediacentre/factsheets/fs297/es/>

## 8. Anexos

### 8.1.Anexo I: Muestras procedentes de un mismo paciente

| Muestra | Procedencia    | Mutaciones divergentes (EGFR) | Mutaciones comunes (EGFR) |
|---------|----------------|-------------------------------|---------------------------|
| M01.1   | Pulmón         | (-)                           | p.L858R                   |
| M01.2   | Pulmón         | p.T790M                       |                           |
| M03.1   | Pulmón         | (-)                           | p.E746-A750del            |
| M03.2   | Mtx: Peritoneo | p.T790M                       |                           |
| M04.1   | Pulmón         | (-)                           | p.L858R                   |
| M04.2   | Pulmón         | p.T790M                       |                           |
| M11.1   | Pulmón         | (-)                           | p.E709K, p.G719C          |
| M11.2   | Pleura         | (-)                           |                           |

## 8.2.Anexo II: chips empleados para la secuenciación mediante PGM

| Chip | Muestras   | Porcentaje de pocillos ocupados (%) | Reads Totales | Reads Validos | Porcentaje de lecturas validas (%) | Media de bases por Read (bp) |
|------|--|-------------------------------------|---------------|---------------|------------------------------------|------------------------------|
| 1    | M01.1, M01.2, M02, M03.1, M03.2, M04.1, M04.2, M05 | 69                                  | 7841057       | 5121500       | 65,31                              | 110                          |
| 2    | M10, M11.1, M11.2, M12, M13, M25                   | 67                                  | 7536289       | 2673687       | 35,48                              | 114                          |
| 3    | M15, M16, M17                                      | 67                                  | 7517968       | 4008658       | 53,32                              | 115                          |
| 4    | M18, M19, M20, M21, M22                            | 78                                  | 8775044       | 3579861       | 40,79                              | 110                          |
| 5    | M06, M07, M08, M23, M24                            | 77                                  | 8639955       | 3121595       | 36,13                              | 111                          |
| 6    | M14, M26, M27, M28                                 | 49                                  | 7646706       | 3664597       | 47,92                              | 113                          |
| 7    | M09, M29, M30, M32                                 | 50                                  | 7944207       | 3911866       | 49,24                              | 112                          |
| 8    | M31, M33, M34, M35, M36                            | 75                                  | 8468926       | 4014029       | 47,39                              | 112                          |

### 8.3.Anexo III: Mutaciones detectadas mediante secuenciación en el PGM

(NV – No valorable)

| Muestra | Procedencia       | Chip | Nº Lecturas | Gen    | Mutación       | Frecuencia alélica (%) |
|---------|-------------------|------|-------------|--------|----------------|------------------------|
| M01.1   | Pulmón            | 1    | 519,807     | EGFR   | p.L858R        | 20,20                  |
|         |                   |      |             | PIK3CA | p.G1049R       | 19,35                  |
| M01.2   | Pulmón            | 1    | 554,289     | EGFR   | p.L858R        | 25,60                  |
|         |                   |      |             |        | p.T790M        | 12,37                  |
|         |                   |      |             | PIK3CA | p.G1049R       | 26,60                  |
| M02     | Pulmón            | 1    | 551,861     | EGFR   | p.T751A        | 58,28                  |
| M03.1   | Pulmón            | 1    | 518,364     | EGFR   | p.E746-A750del | 3,03                   |
|         |                   |      |             | TP53   | p.P72R         | 96,99                  |
| M03.2   | Mtx:<br>Peritoneo | 1    | 6,351       | NV     | NV             | (-)                    |
| M04.1   | Pulmón            | 1    | 402,052     | EGFR   | p.L858R        | 11,00                  |
|         |                   |      |             | TP53   | p.P72R         | 95,23                  |
| M04.2   | Pulmón            | 1    | 503,515     | EGFR   | p.L858R        | 50,49                  |
|         |                   |      |             |        | p.T790M        | 9,22                   |
|         |                   |      |             | TP53   | p.P72R         | 97,00                  |
|         |                   |      |             |        | p.G105V        | 63,22                  |
|         |                   |      |             |        |                |                        |
| M05     | Pulmón            | 1    | 524,024     | PIK3CA | p.H1047L       | 76,41                  |
|         |                   |      |             | TP53   | p.P72R         | 95,60                  |
|         |                   |      |             |        | p.Y220C        | 66,34                  |
| M06     | Pulmón            | 5    | 366,069     | PTEN   | c.635-1G>C     | 28,34                  |
|         |                   |      |             | TP53   | p.P72R         | 84,10                  |
|         |                   |      |             |        | p.R158L        | 69,93                  |
| M07     | Pulmón            | 5    | 274,085     | KRAS   | p.G12V         | 39,38                  |
|         |                   |      |             | AKT1   | p.D46E         | 42,31                  |
|         |                   |      |             | TP53   | p.P72R         | 96,02                  |
|         |                   |      |             |        | p.Y163C        | 49,69                  |
| M08     | Pulmón            | 5    | 292,515     | KRAS   | p.Q61H         | 39,23                  |
|         |                   |      |             | TP53   | p.P72R         | 44,79                  |
| M09     | Ganglio           | 7    | 298,611     | KRAS   | p.G12C         | 23,67                  |
|         |                   |      |             | TP53   | p.P72R         | 97,77                  |
| M10     | Pulmón            | 2    | 374,272     | KRAS   | p.G13E         | 31,89                  |
|         |                   |      |             | TP53   | p.G154fs       | 20,37                  |
| M11.1   | Pulmón            | 2    | 47,478      | NV     | NV             | (-)                    |
| M11.2   | Pulmón            | 2    | 330,619     | EGFR   | p.E709K        | 62,35                  |
|         |                   |      |             |        | p.G719C        | 63,03                  |
|         |                   |      |             | CTNNB1 | p.S37C         | 44,51                  |
|         |                   |      |             | TP53   | p.P72R         | 53,80                  |
| M12     | Pulmón            | 2    | 303,224     | EGFR   | p.L858R        | 6,81                   |
|         |                   |      |             | TP53   | p.P72R         | 51,52                  |
|         |                   |      |             |        | p.L194F        | 8,59                   |

|            |                      |   |         |        |                |        |
|------------|----------------------|---|---------|--------|----------------|--------|
| <b>M13</b> | Pulmón               | 2 | 404,913 | KRAS   | p.G13D         | 17,05  |
|            |                      |   |         | TP53   | p.P72R         | 42,47  |
| <b>M14</b> | Pulmón               | 6 | 288,119 | TP53   | p.V157F        | 25,13  |
| <b>M15</b> | Pulmón               | 3 | 561,364 | TP53   | p.P72R         | 57,11  |
|            |                      |   |         | BRAF   | p.V600E        | 3,20   |
| <b>M16</b> | Pulmón               | 3 | 613,012 | TP53   | p.P72R         | 97,84  |
|            |                      |   |         | KRAS   | p.Q61H         | 13,14  |
|            |                      |   |         | BRAF   | p.V600E        | 5,13   |
| <b>M17</b> | Pulmón               | 3 | 581,215 | TP53   | p.P72R         | 97,70  |
|            |                      |   |         | KRAS   | p.G12V         | 12,29  |
| <b>M18</b> | Pulmón               | 4 | 431,259 | TP53   | p.P72R         | 97,44  |
|            |                      |   |         | EGFR   | c.2361G>C      | 100,00 |
|            |                      |   |         | BRAF   | p.G469V        | 26,23  |
| <b>M19</b> | Pulmón               | 4 | 377,753 | TP53   | p.P72R         | 38,16  |
|            |                      |   |         | KRAS   | p.G12C         | 14,96  |
| <b>M20</b> | Pulmón               | 4 | 389,443 | TP53   | p.P72R         | 100,00 |
|            |                      |   |         | MET    | p.M362T        | 52,90  |
|            |                      |   |         | KRAS   | p.G12V         | 27,50  |
|            |                      |   |         | STK11  | p.P281fs       | 16,20  |
|            |                      |   |         | TP53   | p.P72R         | 97,74  |
| <b>M21</b> | Pulmón               | 4 | 359,729 | TP53   | p.R248L        | 28,60  |
|            |                      |   |         | KRAS   | p.G12V         | 25,00  |
| <b>M22</b> | Mtx: cerebro         | 4 | 369,296 | TP53   | p.P72R         | 97,75  |
|            |                      |   |         | MET    | p.E168D        | 50,40  |
|            |                      |   |         | TP53   | p.P72R         | 35,10  |
| <b>M23</b> | Mtx: tejido linfoide | 5 | 331,411 | TP53   | p.G154V        | 70,74  |
|            |                      |   |         | KRAS   | p.G12C         | 59,69  |
|            |                      |   |         | TP53   | p.P72R         | 80,19  |
| <b>M24</b> | Pulmón               | 5 | 327,735 | TP53   | p.R273L        | 52,14  |
|            |                      |   |         | KRAS   | p.G12C         | 8,63   |
| <b>M25</b> | Pleura               | 2 | 362,708 | TP53   | p.P72R         | 100,00 |
|            |                      |   |         | DDR2   | p.R260L        | 44,99  |
| <b>M26</b> | Pulmón               | 6 | 326,226 | TP53   | p.P72R         | 50,65  |
|            |                      |   |         | KRAS   | p.G12V         | 6,00   |
| <b>M27</b> | Pulmón               | 6 | 415,556 | TP53   | p.P72R         | 55,02  |
|            |                      |   |         | KRAS   | p.G12V         | 22,40  |
| <b>M28</b> | Pulmón               | 6 | 335,933 | TP53   | p.P72R         | 50,62  |
|            |                      |   |         | ERBB2  | p.I767M        | 9,90   |
|            |                      |   |         | DDR2   | p.M441I        | 51,77  |
|            |                      |   |         | PIK3CA | p.E542Q        | 7,69   |
|            |                      |   |         | TP53   | p.P72R         | 98,07  |
| <b>M29</b> | Mtx: Hígado          | 7 | 448,994 | TP53   | p.V157F        | 9,82   |
|            |                      |   |         | EGFR   | p.A750P        | 4,83   |
|            |                      |   |         | MET    | p.E168D        | 30,04  |
|            |                      |   |         | KRAS   | p.G12C         | 79,69  |
| <b>M30</b> | Pulmón               | 7 | 443,395 | TP53   | p.P72R         | 97,16  |
|            |                      |   |         | KRAS   | p.G13E         | 45,12  |
| <b>M31</b> | Pulmón               | 8 | 348,584 | TP53   | p.P72R         | 98,01  |
|            |                      |   |         | EGFR   | p.E746-A750del | 42,18  |

|            |             |   |         |      |                    |        |
|------------|-------------|---|---------|------|--------------------|--------|
|            |             |   |         |      | p.T790M            | 23,44  |
|            |             |   |         | TP53 | p.P72R             | 100,00 |
| <b>M32</b> | Pulmón      | 7 | 300,795 | EGFR | p.L858R            | 15,32  |
|            |             |   |         | TP53 | p.P72R             | 100,00 |
| <b>M33</b> | Pulmón      | 8 | 540,709 | KRAS | p.G12C             | 16,03  |
| <b>M34</b> | Pulmón      | 8 | 333,021 | KRAS | p.G12C             | 50,90  |
|            |             |   |         | TP53 | p.P72R             | 100,00 |
|            |             |   |         |      | p.R249S            | 36,00  |
| <b>M35</b> | Pulmón      | 8 | 445,948 | KRAS | p.Gly12Cys         | 28,34  |
|            |             |   |         |      | p.Thr58Ile         | 2,45   |
|            |             |   |         | TP53 | p.Pro72Arg         | 100,00 |
| <b>M36</b> | Mtx: Hígado | 8 | 347,891 | EGFR | p.L747-P753delinsS | 51,00  |
|            |             |   |         |      | p.T790M            | 12,50  |
|            |             |   |         |      | p.C797S            | 12,50  |
|            |             |   |         | MET  | p.N375S            | 77,00  |
|            |             |   |         | TP53 | p.P72R             | 78,55  |

#### 8.4.Anexo IV: Resultados de la secuenciación de los transcritos

(NV – No valorable)

| Muestra | Chip | Mapped Reads | Traslocación identificada/validada |
|---------|------|--------------|------------------------------------|
| M01.1   | 1    | 154,903      |                                    |
| M01.2   | 1    | 157,106      |                                    |
| M02     | 1    | 3,499        | NV                                 |
| M03.1   | 1    | 121,073      |                                    |
| M03.2   | 1    | 107,044      |                                    |
| M04.1   | 1    | 29,839       |                                    |
| M04.2   | 1    | 124,945      |                                    |
| M05     | 1    | 153,58       |                                    |
| M06     | 5    | 154,201      |                                    |
| M07     | 5    | 120,828      |                                    |
| M08     | 5    | 200,810      |                                    |
| M09     | 7    | 177,102      |                                    |
| M10     | 2    | 218,315      |                                    |
| M11.1   | 2    | 221,595      |                                    |
| M11.2   | 2    | 147,269      |                                    |
| M12     | 2    | 291,415      |                                    |
| M13     | 2    | 168,582      |                                    |
| M14     | 6    | 144,357      | EML4(13) - ALK(20)                 |
| M15     | 3    | 113,784      |                                    |
| M16     | 3    | 100,978      |                                    |
| M17     | 3    | 122,335      |                                    |
| M18     | 4    | 116,243      |                                    |
| M19     | 4    | 65,368       |                                    |
| M20     | 4    | 79,330       |                                    |
| M21     | 4    | 130,086      |                                    |
| M22     | 4    | 104,748      |                                    |
| M23     | 5    | 75,829       |                                    |
| M24     | 5    | 159          | NV                                 |
| M25     | 2    | 249,939      | EML4(6) - ALK(20)                  |
| M26     | 6    | 102,082      |                                    |
| M27     | 6    | 87,849       |                                    |
| M28     | 6    | 53,343       |                                    |
| M29     | 7    | 88,457       |                                    |
| M30     | 7    | 79,743       |                                    |
| M31     | 8    | 67,713       |                                    |
| M32     | 7    | 49,351       |                                    |
| M33     | 8    | 114,948      |                                    |
| M34     | 8    | 112,940      |                                    |
| M35     | 8    | 163,197      |                                    |
| M36     | 8    | 112,394      |                                    |